2014

# Integrated Remote Sensing and Forecasting of Regional Terrestrial Precipitation with Global Nonlinear and Nonstationary Teleconnection Signals Using Wavelet Analysis

Lee Mullon
*University of Central Florida*

# INTEGRATED REMOTE SENSING AND FORECASTING OF REGIONAL TERRESTRIAL PRECIPITATION WITH GLOBAL NONLINEAR AND NONSTATIONARY TELECONNECTION SIGNALS USING WAVELET ANALYSIS

by

LEE GRAY MULLON
B.S. University of Central Florida, 2005

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science in Civil Engineering, Water Resources Engineering Track
in the Department of Civil, Environmental, and Construction Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Fall Term
2014

Major Professor: Ni-Bin Chang

## ABSTRACT

Global sea surface temperature (SST) anomalies have a demonstrable effect on terrestrial climate dynamics throughout the continental U.S. SST variations have been correlated with greenness (vegetation densities) and precipitation via ocean-atmospheric interactions known as climate teleconnections. Prior research has demonstrated that teleconnections can be used for climate prediction across a wide region at sub-continental scales. Yet these studies tend to have large uncertainties in estimates by utilizing simple linear analyses to examine chaotic teleconnection relationships. Still, non-stationary signals exist, making teleconnection identification difficult at the local scale. Part 1 of this research establishes short-term (10-year), linear and non-stationary teleconnection signals between SST at the North Atlantic and North Pacific oceans and terrestrial responses of greenness and precipitation along multiple pristine sites in the northeastern U.S., including (1) White Mountain National Forest – Pemigewasset Wilderness, (2) Green Mountain National Forest – Lye Brook Wilderness and (3) Adirondack State Park – Siamese Ponds Wilderness. Each site was selected to avoid anthropogenic influences that may otherwise mask climate teleconnection signals. Lagged pixel-wise linear teleconnection patterns across anomalous datasets found significant correlation regions between SST and the terrestrial sites. Non-stationary signals also exhibit salient co-variations at biennial and triennial frequencies between terrestrial responses and SST anomalies across oceanic regions in agreement with the El Nino Southern Oscillation (ENSO) and North Atlantic Oscillation (NAO) signals. Multiple regression analysis of the combined ocean indices explained up to 50% of the greenness and 42% of the precipitation in

the study sites. The identified short-term teleconnection signals improve the understanding and projection of climate change impacts at local scales, as well as harness the interannual periodicity information for future climate projections.

Part 2 of this research paper builds upon the earlier short-term study by exploring a long-term (30-year) teleconnection signal investigation between SST at the North Atlantic and Pacific oceans and the precipitation within Adirondack State Park in upstate New York. Non-traditional teleconnection signals are identified using wavelet decomposition and teleconnection mapping specific to the Adirondack region. Unique SST indices are extracted and used as input variables in an artificial neural network (ANN) prediction model. The results show the importance of considering non-leading teleconnection patterns as well as the known teleconnection patterns. Additionally, the effects of the Pacific Ocean SST or the Atlantic Ocean SST on terrestrial precipitation in the study region were compared with each other to deepen the insight of sea-land interactions. Results demonstrate reasonable prediction skill at forecasting precipitation trends with a lead time of one month, with r values of 0.6. The results are compared against a statistical downscaling approach using the HadCM3 global circulation model output data and the SDSM statistical downscaling software, which demonstrate less predictive skill at forecasting precipitation within the Adirondacks.

For Christine and Isabel

## ACKNOWLEDGMENTS

The author wishes to thank his Thesis Committee, Drs. Martin Wanielista, Dingbao Wang, and Chair Ni-Bin Chang.

## TABLE OF CONTENTS

## LIST OF FIGURES

# LIST OF TABLES

## LIST OF ACRONYMS (or) ABBREVIATIONS

| | |
|---|---|
| ANN | Artificial Neural Network |
| AO | Arctic Oscillation |
| ARMA | Autoregressive Moving Average |
| AVHRR | Advanced Very High Resolution Radar |
| DJF | December-January-February |
| CPC | Climate Prediction Center |
| CWT | Continuous Wavelet Transform |
| ENSO | El Nino Southern Oscillation |
| EOF | Empirical Orthogonal Function |
| EVI | Enhanced Vegetation Index |
| GCM | General Circulation Models |
| GHRSST | Group for High Resolution Sea Surface Temperature |
| GMNF | Green Mountain National Forest |
| IPCC | Intergovernmental Panel on Climate Change |
| JJA | June-July-August |
| LBW | Lye Brook Wilderness |
| MAM | March-April-May |
| MOC | Meridional Overturning Circulation |
| NAO | North Atlantic Oscillation |
| NDVI | Normalized Difference Vegetation Index |

| | |
|---|---|
| NOAA | National Oceanic and Atmospheric Administration |
| PNA | Pacific North American Pattern |
| PDO | Pacific Decadal Oscillation |
| PW | Pemigewasset Wilderness |
| SAWP | Scale-Averaged Wavelet Power |
| SDSM | Statistical Downscaling Model |
| SLP | Sea Level Pressure |
| SON | September-October-November |
| SOI | Southern Oscillation Index |
| SPW | Siamese Ponds Wilderness |
| SST | Sea Surface Temperature |
| TPC | Teleconnection-based Principal Components (based on SST extracted regions) |
| USDA | United States Department of Agriculture |
| WMNF | White Mountain National Forest |
| WPC | Wavelet Principal Component (also Wavelet Empirical Orthogonal Function EOF) |
| WTC | Wavelet Coherency Transform |
| XWT | Cross Wavelet Transform |

# CHAPTER 1 : INTRODUCTION

Climate change and variability estimates from the Intergovernmental Panel on Climate Change (IPCC) in 2007 forecast an increase in North American temperature and precipitation during the 21st century (Meehl & Stocker, 2007). These fundamental natural processes are a result of the complex and chaotic interactions of the Earth system, which can be described as a collection of subsystems of interconnected parts or spheres, including the atmosphere, biosphere, hydrosphere and geosphere, each of which is open and allows the flow of mass and energy from one to another (Rial et al., 2004).

Abrupt changes in temperature and precipitation can have a profound effect on the natural environment. Global vegetation density, or greenness, is highly correlated with temperature and precipitation (Los et al., 2001), and is susceptible to future climate variability. Yet it is unclear how teleconnection signal propagations could affect local temperature and precipitation patterns and subsequent ecosystem response, such as vegetation greenness, across different locales of North America. These IPCC forecasts trigger a renewed interest to study the response of precipitation patterns and terrestrial ecosystems in association with some hydrometeorological forcing processes of circumglobal teleconnection in the context of climate change. Such circumglobal teleconnection in the context of this paper refers to climate anomalies being related to each other at far distances (often thousands of kilometers). Understanding these relationships can yield new insight into forecasting the effects of climate change, as they relate to environmental time series data, such as precipitation or greenness.

New tools in remote sensing and spatial statistical analysis now permit deepened investigation of the mechanistic links between changing teleconnection patterns and their residual effect or memory on local terrestrial environments (David B. Enfield et al., 2001; Hodson et al., 2009; Keener et al., 2010; Raible, 2003; Sutton and Hodson, 2007). These mechanistic links can then provide a basis to estimate future impact on local precipitation and greenness in association with teleconnection patterns through various hydrometeorological forcing processes (Cai et al., 2009; Keener et al., 2010).

### 1.1 Teleconnections

Strong oceanic, hydrometeorological and ecological phenomena are physically linked across the globe by atmospheric circulation through some excitation mechanisms (Buermann, 2003; Franzke and Woollings, 2011; Hubeny et al., 2011; Johnstone, 2010; Joseph and Nigam, 2006; Sutton and Hodson, 2005). Climate teleconnections are often categorized in known oceanic–atmospheric circulation patterns, such as the El Nino Southern Oscillation (ENSO), North Atlantic Oscillation (NAO), Arctic Oscillation (AO) and the Pacific North American Pattern (PNA). These climatic patterns can be linked with time-series observations based on a multivariate array of oceanic (e.g., sea surface temperature, SST) and atmospheric (e.g., sea level pressure, SLP) variables. Teleconnection signals can also be found between univariate oceanic variables and inland terrestrial temperature, precipitation or greenness (Huber and Fensholt, 2011).

Of particular interest in teleconnection research is the exploration of specific periodicity of these climate signals over which multiple terrestrial phenomena can be linked with each

other. Long-term multidecadal teleconnections have been observed and modeled over North America (David B Enfield et al., 2001; Franzke and Woollings, 2011; Hodson et al., 2009; Sutton and Hodson, 2007). These studies demonstrated low-frequency covariability in periods of 80 years between Atlantic and Pacific SST-related climate indices and the variability of temperature and precipitation in North America. Higher-frequency teleconnections ranging from days to biennial and triennial periods also have been observed (Athanasiadis and Ambaum, 2009). Johnstone (2011) documented more than 30 global quasi-biennial (25–26 months) temperature and precipitation oscillations that were discovered during the past 40 years, with six existing cases over the United States; yet the study sites involved with coupled natural forcing and anthropogenic disturbance (such as urban areas) made teleconnection signals unidentifiable.

The identification of distinct climatic teleconnection signals has long been used for specific forecasting of various hydrologic events. In traditional teleconnection identification, a linear correlation is often assumed between anomalous SST and SLP across the Pacific Ocean (Montroy, 1997) for projections of many local and regional terrestrial responses throughout the northern hemisphere, including the number and location of wildfires (Dixon et al., 2008), groundwater recharge rates (Holman et al., 2011), streamflow levels and nutrient loading (Keener et al., 2010). More recently, nonlinear and nonstationary tools have been developed to explore how the low- and high-frequency behavior between teleconnections and hydrologic system variables interact, which can lead to the development of robust forecasts in the field of hydrologic time series prediction.

## 1.2 Hydrologic Time Series Prediction

Hydrologic time series prediction has been a successful modeling tool to address the underlying physical mechanism through a holistic approach (Coulibaly and Baldwin, 2005). A problem arises in that the physical system is highly dynamic, exhibiting nonlinearity and nonstationarity, meaning that the characteristics of the system change over time due to external forcings and internal feedback mechanisms. Inputs and outputs of the systems are not proportional, and they are subject to frequent abrupt change, as opposed to slow and gradual, where multiple states of equilibrium are normal (Rial et al., 2004). This is in contrast with traditional assumptions of stationarity in many environmental systems analyses.

Climate, and therefore climatic processes (e.g. precipitation), is generally understood as the best example of nonlinear dynamics. Drivers of the climate system, such as the Atlantic Meridional Overturning Circulation (MOC), have been shown to exhibit multiple states of equilibrium and thresholds that result in periods of abrupt climate change (Marshall et al., 2001). Such traits are characteristics of nonlinear behaviour.

Hydrologic forecasting has historically been performed using environmental time series data, which are essential tools for the planning and management of water resource systems including water quality, water demand, water pricing and hydrometeorological processes (Maier and Dandy, 2000). These modeled time series data have historically relied upon the use of multivariate Auto Regressive Moving Average (ARMA) models (Rao and Krishna, 2009), which have long been considered an acceptable and successful representation of stochastic time series (Young et al., 1991). However, Coulibaly and Baldwin (2005) pointed out that the

fundamental assumption in most empirical or statistical approaches, such as the ARMA models, is stationarity over time, which cannot fully explain the complexities of geophysical data. Such difficulties expressing the geophysical system in traditional models have given rise to the use of artificial neural network (ANN) models for nonlinear, nonstationary data analysis.

### 1.3 Objectives and Organization

The objective of this study was to examine the low- and high-frequency teleconnection signals in association with hydrometeorological patterns between the Northern Atlantic and Pacific oceans simultaneously at pristine terrestrial forest regions of northeastern U.S., with limited anthropogenic disturbance in the study period, and to determine if the discovered statistical relationships could be leveraged in a hydrologic forecasting scheme. In the context of high temporal resolution for hydroclimatic applications, the high-frequency teleconnections are defined here as correlation signals at the intraannual scale, whereas the low-frequency scale is for the interannual scales.

This study is divided into two primary research investigations. In the first investigation, a short-term (10 year) study researches climate teleconnections between SST, precipitation and greenness in three pristine forested sites in the northeastern U.S., including (1) White Mountain National Forest - Pemigewasset Wilderness, (2) Green Mountain National Forest – Lye Brook Wilderness, and (3) Adirondack State Park – Siamese Ponds Wilderness, to gain a statistical insight of relevance region wide. The short-term time scale was chosen based on the availability of greenness Enhanced Vegetation Index (EVI) data, which are at high spatial resolution scales from satellite remote sensing imagery. Integrated remote sensing and wavelet analysis were

employed to detect distinct interannual (2 to 3 years) signals of enhanced variability suggestive of these aforementioned teleconnection patterns into the region of northeast U.S. The teleconnection signal propagation was investigated by using statistical methods, including a pixel-wise correlation between precipitation, greenness and SST, as well as spectral analyses of the time series data using wavelet analysis. The short-term research helps confirm how low- and high-frequency transients through the teleconnections affect terrestrial vegetation cover and precipitation patterns at the three predetermined pristine forested sites.

The second investigation performs a long-term (30 year) study which focuses only on the teleconnection between SST in the northern Atlantic and Pacific and the precipitation at the Adirondack State Park. In addition to a teleconnection investigation, the long-term research also includes an ANN-based precipitation forecasting scheme developed to predict seasonal and monthly precipitation within the Adirondack boundary.

The following research is organized into seven chapters, including Chapter 1, this Introduction; Chapter 2 will present a detailed literature review of teleconnections and hydrologic time series prediction models; Chapter 3 will present the short-term teleconnection investigation and research findings; Chapter 4 will present the long-term teleconnection investigation, forecasting models and research findings; Chapter 5 will present the concluding remarks; and Chapter 6 will present the list of references and citations.

# CHAPTER 2 : LITERATURE REVIEW

## 2.1 Teleconnections in the Northeastern U.S.

Observations during the past 1,000 years indicate the strong role that the Pacific and Atlantic oceans play in the New England area (Hubeny et al., 2011). The correlation signals between the PNA, NAO and ENSO and the New England climate are still predominant in the precipitation and temperature response for the region today. Precipitation amounts over the past century have increased by 10 to 15% and are expected to increase up to 34% during the next century (Beckage et al., 2008; Tang and Beckage, 2010). This increased precipitation has been influenced by positive cycles of PNA, characterized by above-average temperatures in western Canada and the U.S. and below-average temperatures across the south-central and southeastern U.S. (National Weather service Climate Prediction Center, 2012).

ENSO, a coupled ocean-atmospheric system, has a strong influence over the global climate system, including the New England area. Per the NOAA National Climate Data Center (NCDC), the Southern Oscillation Index (SOI) is a standardized index based on the observed sea level pressure differences between Tahiti and Darwin, Australia, which is one measure of the large-scale fluctuations in air pressure occurring between the western and eastern tropical Pacific (i.e., the state of the Southern Oscillation) during El Niño and La Niña episodes (Center, 2012). The negative phase of the SOI represents below-normal air pressure at Tahiti and above-normal air pressure at Darwin. Prolonged periods of negative (positive) SOI values coincide with abnormally warm (cold) ocean waters across the eastern tropical Pacific typical of

El Niño (La Niña) episodes. Linking with a difference in tropical Pacific sea level pressure between Tahiti and Darwin, ENSO has a general oscillation period of 2 to 7 years (National Weather Service Climate Prediction Center, 2012). The warm El Niño phase influences climate across the eastern U.S. by shifting the jet stream southward, bringing cold winter weather, whereas the cold La Niña cycle shifts the jet stream poleward, bringing warmer and milder winter weather to the same region (Greene, 2012). Pacific-Atlantic communication may also exist along the equator, with the tropical Pacific modulating the Hadley cell across Central and South America. Along this teleconnection, Atlantic SST and wind patterns are influenced by ENSO, with ENSO-derived Atlantic SST variations lagging the Pacific patterns by 4–5 months (Enfield and Mayer, 1997) based on various interannual time scales. Thus, New England climate associated with Atlantic-based variability may originate in the Pacific.

The PNA pattern is a broad reflection of the upper-level jet stream, which typically consists of a large ridge and trough pattern across the U.S. The PNA index details the oscillation of this Northern Hemispheric wave (Leathers et al., 1991). The phase of PNA, and therefore the location of the jet stream, strongly influences the storm track and weather patterns in the northeast. Notaro et al. (2006) identified the positive PNA pattern as influencing winter time drought conditions in the northeast, as well as frequent frontal passages through New York during negative PNA and positive NAO cycles, with a zonal upper-level jet generally positioned over New York. PNA is also influenced by ENSO, especially during strong ENSO periods, which modulates the position of the East Asian jet stream. Positive PNA patterns are often observed

during El Niño periods, and negative patterns associated with La Niña periods (National Weather Service Climate Prediction Center, 2012).

Atlantic patterns have a clear effect on the New England climate, particularly temperature. The NAO is the dominant pattern of climate variability in the northern Atlantic and is sometimes considered the Atlantic component of the northern hemispheric annular mode, the AO. NAO is simply the SLP difference between the Icelandic Low and the Azores High (Marshall et al., 2001). Other primary modes of variability in the northern Atlantic include the Tropical Atlantic Variability (TAV), consisting of the fluctuations in tropical Atlantic SST and trade winds along the Intertropical Convergence Zone (ITCZ), and the Atlantic MOC, the Atlantic component of the fluctuating thermohaline circulation. The NAO atmospheric pattern is most dominant in the mid-latitudes, where NAO-derived extratropical forcings excite the TAV by rearranging the the Hadley circulation patterns and ultimately the location of the ITCZ. In addition, by modulating the location and intensity of the sinking branch of the MOC, the NAO effectively controls the primary heat transfer mechanism of the equatorial Atlantic to the northern latitudes. This broad influence over the North Atlantic basin is why NAO can be linked to such wide variations of temperature, precipitation, storm track and ecosystem variability across wide reaches in North America, Europe, Asia and North Africa, and why NAO can be considered a rival teleconnection pattern to ENSO in many regions, including New England. In New England, negative NAO phases (warm, high-pressure periods near Iceland) are marked by colder temperature, while the opposite is true for positive phases. This temperature pattern is a result of a fluctuating jet stream.

During positive NAO phases, Icelandic low-pressure systems and Azores high pressures lock the jet stream in a northeastern zonal path across North America toward Western Europe. During negative NAO phases, higher-pressure systems near Iceland weaken the atmospheric pressure gradient, which causes the jet stream to behave more meriodonal, often moving far southward and bringing polar weather south to the U.S. Because jet stream forcings are common to both ENSO and NAO, periods of both El Niño and negative NAO phases can cause unusually cold weather and storm tracking across New England. SST anomalies in the North Atlantic are linked to precipitation variability throughout the region, although to a lesser extent than temperature (Marshall et al., 2001). Because precipitation and temperature are the primary mechanisms of vegetation dynamics, greenness patterns are expected to correlate directly with regional precipitation and temperature and indirectly with the NAO–SST mode.

The climatic trend and ecosystem response associated with the northeastern U.S. suggest that the climate at the time of European colonization had a strong nonlinear control over broad-scale vegetation patterns (Foster et al., 1998). This correlation signal has been significantly weakened over the past 350 years due to increasing anthropogenic impacts to ecosystems. This signal change must therefore be considered in analyzing the relationships between contemporary climate change and ecosystem dynamics. For this reason, a screening procedure for site selection was developed in this study to isolate pristine, heavily vegetated environments that nullified, to the greatest extent possible, the influence of anthropogenic "noise" on the climate teleconnection signal. This procedure allowed the application of wavelet

multi-resolution spectral analysis to retrieve the nonlinear and non-stationary teleconnection patterns.

## 2.2 Artificial Neural Networks

Artificial Neural Networks were conceived in mid-20$^{th}$ century (McCulloch and Pitts, 1943) to approximate the architecture of the biological neural network of the human brain. The highly distributed information processing of ANNs enables them to identify underlying relationships and interactions in highly complex nonlinear data, and is well suited for nonlinear modelling, pattern recognition, classification and control. Over time, the capabilities of ANNs became increasingly popular, and by the 1990s were beginning to be used in the field of water resources. Several hydrologic disciplines requiring multivariate forecasts have successfully implemented ANN modeling, including rainfall-runoff estimation (Minns and Hall, 1996), stream flow forecasting and water quality (Keener et al., 2010), ground-water modeling (Holman et al., 2011), precipitation forecasting (Abbot and Marohasy, 2012), water management policy and reservoir operations (ASCE, 2000). Some of the key advantages that enabled ANNs success over traditional ARMA models include input data that do not require Gaussian distribution, seasonal variability allowance, strong performance with limited data, nonlinear capabilities and the ability to approximate the underlying relationships within the data without requiring a physical model or fundamental understanding of the system (Maier, 1997).

Input data of ANNs can vary widely depending upon the goals of the model. All models, however, attempt to limit the input (predictor) variables to only those that can improve the model results, because too many variables, or inconsequential ones, can cause unnecessary

complexity, expense and poor performance. Traditional predictor parameters include the historic values of the time series in which the model is based (e.g. streamflow), as well as other exogenous variables the modeler expects are related to the modeled phenonmena (e.g. precipitation or evapotranspiration).

Recently ANN models have found climate teleconnection patterns as useful predictor variables for hydrologic forecasting. These teleconnection patterns take the form of published climate indices of variable oceanic-atmospheric patterns across the globe. The ENSO pattern, for example, is an important coupled oceanic-atmospheric phenomena with global significance to the Earth's climate, with strong manifestations in temperature and precipitation worldwide (IPCC, 2007). Abbot and Marohasy (2012) found a strong performance enhancement to Australian precipitation ANN forecasting when combining multiple climate teleconnection indices, including the SOI, Pacific Decadal Oscillation (PDO) and the Niño 3.4 index.

ANNs are inherently reliant on the basic time series predictor data of the model. Early ANN-based hydrologic studies reported limited forecasting success, which has been attributed to weaknesses in the the time series inputs. To solve this, (Wang and Ding, 2003) proposed to pre-treat the time series data by decomposing them into their fundamental signals, using wavelets, in a new hybrid ANN model called Wavelet Neural Network (WNN). Wavelet decomposition provides a mathematical process for distilling a signal into multiple levels of details, while also extracting local information of the time series (Rao and Krishna, 2009), providing a more robust representation of the time series data. WNNs also solve a common problem in ANN forecasting, in that multivariate hydrologic time series data are routinely

combined from many sources of varying frequencies into a single model, such as evaportransporation and streamflow data (Anctil and Tape, 2004). This can have a conflicting effect on the ANN, causing the model to diverge from the solution. By extracting the underlying frequency behavior of these datasets, the WNN models can more successfully interpret the interrelationships.

WNN forecasting has been successfully applied in a number of hydrologic processes, including rainfall runoff, groundwater level and precipitation prediction. These studies have primarily modeled point source areas, such as stream and rain gauge data, to validate forecast skill. Such models are useful at the river or gauge location, or to make reasonable prediction about continental-scale climate patterns, but do not provide the spatial resolution or distribution necessary to help water managers at the local, watershed scale.

Mwale et al. (2004) used a WNN-based model to forecast seasonal precipitation across a gridded array of rainfall data at a continental scale. With a novel approach, Mwale et al. (2004) applied wavelet principal component analysis (WPCA) to individual scale-averaged wavelet power (SAWP) datasets to identify homogeneous zones of rainfall variability and predictability in East Africa. Interestingly, Mwale et al. (2004) did not rely on traditional teleconnection regions, such as ENSO and NAO, to include in the predictive modeling, and instead relied on objectively-identified non-traditional teleconnection regions to obtain predictor data. The research discussed herein builds upon these earlier ANN modeling efforts to leverage existing teleconnection signals to develop more robust predictive models for hydrologic forecasting.

# CHAPTER 3 : SHORT-TERM TELECONNECTION IDENTIFICATION

## 3.1 Methodology

The analytical framework for this study (Figure 3-1) emphasizes the screening and selection of three study sites in relatively pristine forestland. The screening criteria rejected areas with identified forest fires, logging, residential development, hurricane impacts, landslides, debris flows or other relevant natural and anthropogenic impacts. Synchronous remote sensing imageries collected by satellite sensors and corresponding ground-based radar stations were then grouped to meet the requirement of consistent records in SST, precipitation and vegetation cover over the study period. All remote sensing datasets were processed in ArcGIS. To identify the correlation between the SST anomalies and response in terrestrial ecosystem, statistical analysis was performed using MATLAB, with code developed by Grinsted et al. (2004). Ocean indices were then extracted where the high- or low-frequency signal was dominant. Index-based oceanic regions were then identified to detect common low-frequency signals between greenness/SST and precipitation/SST via wavelet analysis.

The following scientific questions were explored in the short-term investigation: (1) are high- and low-frequency signals present between SST anomalies and greenness and precipitation; (2) does precipitation explain a significant portion of greenness variability under the long-standing teleconnection signals; (3) do short-term, localized regions reflect similar teleconnection patterns as those in long-term known teleconnection regions; (4) can combined ocean regions aid in the predictability of local greenness and precipitation variability; and (5) is

there a residual or memory effect in terms of the SST anomalies over the time horizon on terrestrial responses of precipitation and greenness variation.

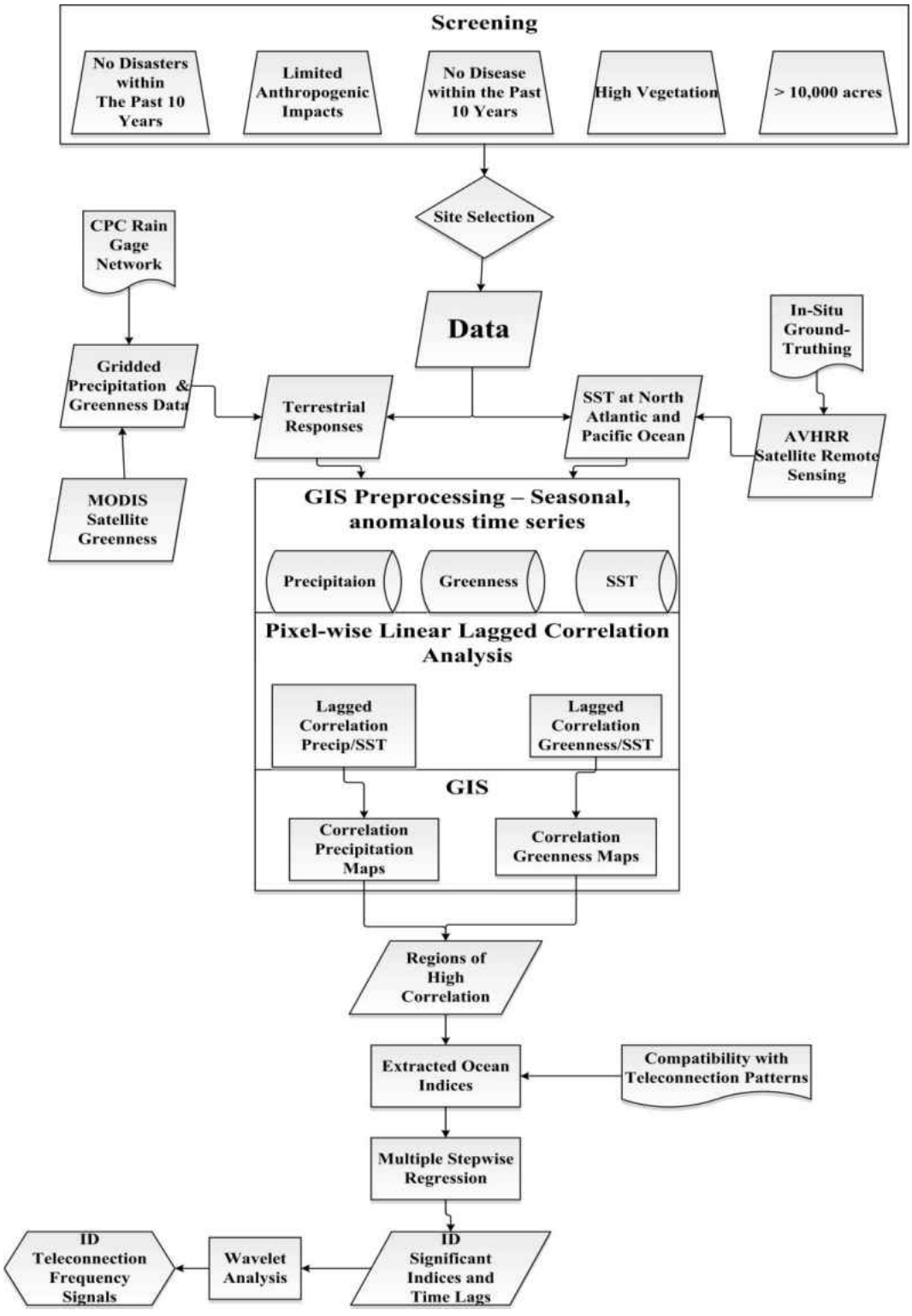


Figure 3-1: Analytical framework for the short-term investigation.

### 3.2 Three Pristine Sites of Little Anthropogenic Influence

Wilderness sites were chosen as the pristine and natural lands for this study, and are defined as "an area where the earth and its community of life are untrammeled by man (*The Wilderness Act*, 1964)." The designation by U.S. Wilderness Act of 1964 restricts commercial activities, permanent roadway, motorized equipment, structure or installation of facilities within designated wilderness areas. The three terrestrial study sites are (1) the Pemigewasset Wilderness within the White Mountain National Forest (WMNF) in New Hampshire; (2) the Lye Brook Wilderness within the Green Mountain National Forest (GMNF) in Vermont; and (3) the Siamese Ponds Wilderness (SPW) within the Adirondack State Park in New York (Figure 3-2). These sites are all highly vegetated areas >10,000 acres (4,046 ha), where climate teleconnection signals with North Atlantic and Pacific SST mode can potentially be greatest. Unlike similar studies (Cho et al., 2010; Holman et al., 2011; Huber and Fensholt, 2011) conducted over vast areas or in urban environments, these three study sites have experienced little to no anthropogenic impact (logging, prescribed burning or mining) or natural disasters (hurricanes or wildfires) within the past three decades. This screening process limited teleconnection signal distortion due to manmade interactions and extreme weather.

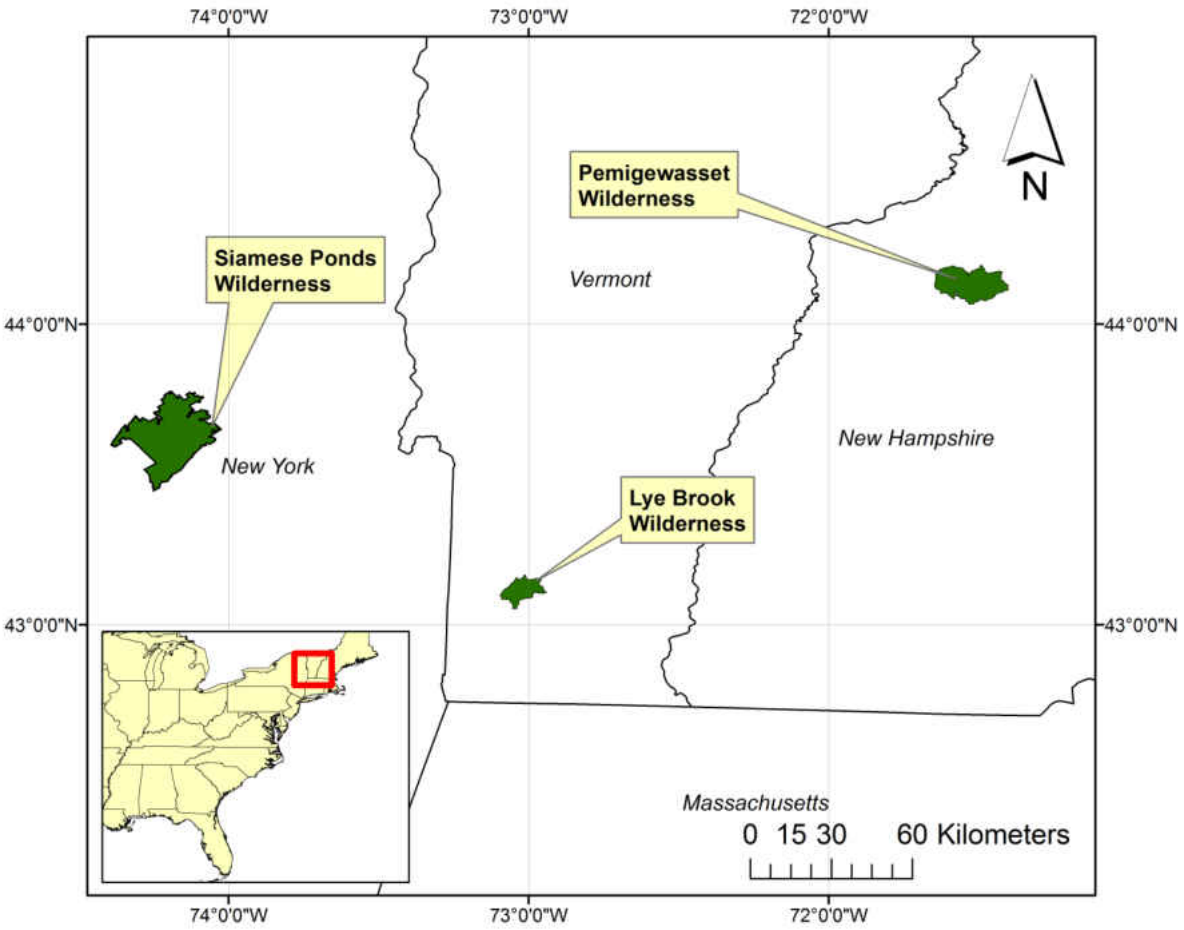


Figure 3-2: Terrestrial study regions including Lye Brook Wilderness, Pemigewasset Wilderness and Siamese Ponds Wilderness.

### 3.2.1 White Mountain National Forest – Pemigewasset Wilderness

The WMNF is located in northern New Hampshire and extends partly into southern Maine, covering an area of 796,700 acres (322,420 ha). It hosts numerous mountain peaks, including 48 that exceed 1,200 m (3,937 ft) and a large alpine zone in the east. Dominant vegetation is softwoods and northern hardwood forests. Nearly 115,000 acres (46,540 ha) of wilderness exist in five separate areas of the forest, including Great Gulf, Presidential

Range/Dry River, Pemigewasset, Sandwich Range and Caribou-Speckled Mountain (USDA, 2008).

The 45,000-acre (18,211 ha) Pemigewasset Wilderness (PW), officially established in the 1984 New Hampshire Wilderness Act, is located within the WMNF. Extensive logging during the railroad era in the 19$^{th}$ and early 20$^{th}$ centuries led to major wildfires in the region. Currently, wildfire is not a common natural disturbance within the wilderness; no significant wildfire has occurred within the past 100 years. Elevations within the PW range from 387 m (1,270 ft) to 1600 m (5,249 ft). The PW is managed only for nonmotorized recreational use (hiking and backpacking).

### 3.2.2 Green Mountain National Forest – Lye Brook Wilderness

The GMNF is located in southwestern and central Vermont, totaling more than 400,000 acres (161,878 ha). Like the WMNF, Green Mountain is characterized by rugged mountain peaks with large northern hardwoods and softwood forests. Early colonial land practices primarily consisted of subsistence farming, grazing and orchard operations. Later, more intense industries including mining and logging reshaped the landscape. Continued degradation of the land, as well as a major flood in 1927, led to legislative actions creating the GMNF in 1932.

Lye Brook Wilderness (LBW) located in southwest GMNF totals 17,841 acres and ranges in elevation between 242 m (794 ft) and 878 m (2,801 ft). Over the past 80 years, the GMNF has been assembled parcel by parcel by land acquisition from private landowners, and it is now the largest contiguous expanse of public land in Vermont. These private lands historically were heavily logged and grazed, with much of their natural resources removed or degraded. Today,

the forest is recognized as an area of significant ecological diversity, including at least 69 unique natural communities ranging from small marsh wetlands to stands of hemlock. Open wetlands, rocky outcrops and cliffs compromise 7,000 acres or 2,833 ha (2% of the forest). Approximately 100,000 acres (40,469 ha) are used for logging and other industry, although not within the LBW. Natural disasters, including wildfire and insect disease, are a noted problem to the forest, including 282 acres (114 ha) of land burned in 2009 alone. The Asian longhorn beetle and the emerald ash borer are two pests that threaten the forest's future (USDA, 2009); although they were not considered significant for the purposes of this study.

### 3.2.3 Adirondack State Park – Siamese Ponds Wilderness

SPW, 114,010 acres (46,139 ha) of forest preserve land, is one of the largest wilderness areas in the Adirondack Park, located in the park's south-central portion within the counties of Warren and Hamilton (New York State Department of Environmental Conservation, 2005). SPW is considered part of the Adirondack highlands, with elevations ranging from 387 m (1,270 ft) to 1,052 m (3,451 ft). In the early 19$^{th}$ century, the forests of the SPW, as in the rest of Adirondack Park, were heavily logged by the timber industry and clear-cut for expanding farms in the region. Several open pit mines were established within SPW; only the garnet mines in Ruby Mountain are currently in operation. After much of the SPW was logged for timber, the land reverted to state ownership in the late 19$^{th}$ century through tax sales. More than 80% of SPW was state-owned by 1910.

Slash and debris from the logging era accumulated, and severe drought in the early 20$^{th}$ century brought the great fires that burned thousands of acres within the SPW. Subsequent

natural disasters included the blow-down of 1950, caused by hurricane-force winds that damaged large areas of the western SPW, and major disease outbursts such as beech bark disease have affected the hardwood in large areas since the 1960s.

### 3.2.4 SST Delineation in North Atlantic and Pacific Oceans

The areas marked in both the Northern Atlantic and Pacific oceans (Figure 3-3) were selected to examine SST anomalies and climate signal teleconnection to the three terrestrial sites.

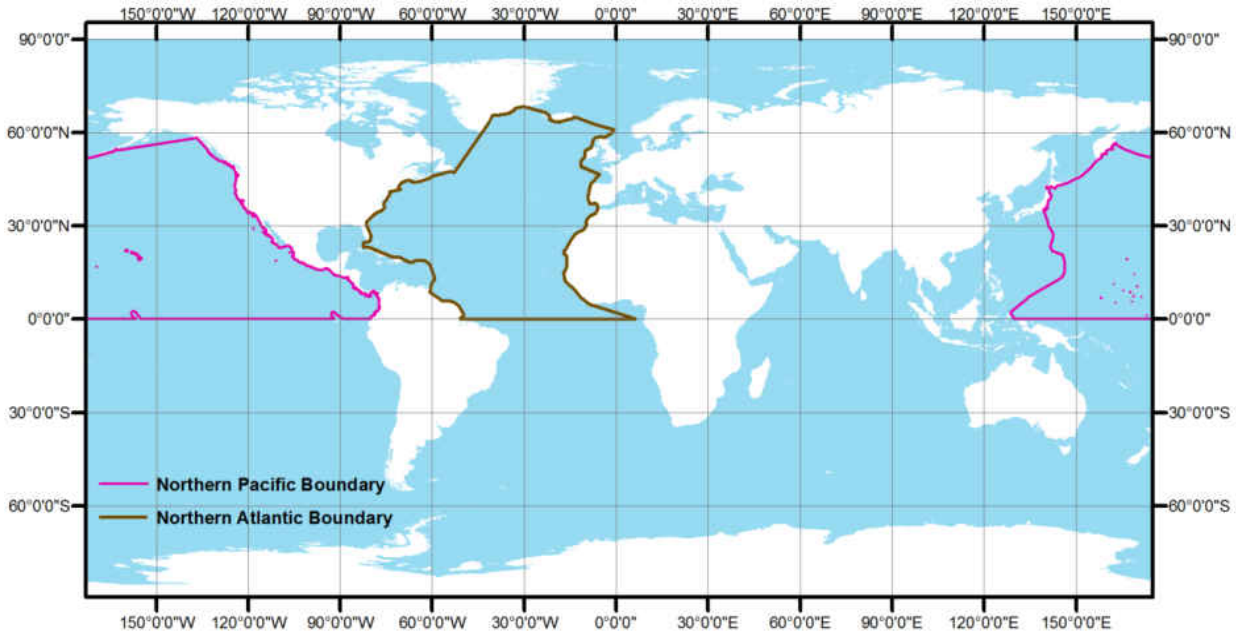


Figure 3-3: Northern Atlantic and Pacific boundaries in this study. The areas marked in both the Northern Atlantic and Pacific oceans were selected to examine SST anomalies and climate signal teleconnection to the three terrestrial sites.

## 3.3 Remote Sensing Data Preparation

Table 3-1: Summary of Remote Sensing Imagery and Instruments.

| Remote Sensing Sensors | MODIS Terra Enhanced Veg. Index (EVI) | AVHRR Pathfinder GHRSST Global Level 4 (SST) | NEXRAD Level 4 Data (Univ. of Okla) (Precipitation) |
|---|---|---|---|
| Data | Detects vegetation density (greenness) | Detects Sea Surface Temperature (SST) | Detects precipitation |
| Temporal Resolution | Monthly data product from daily overpass images | Daily worldwide SST images | Daily precipitation totals |
| Spatial Resolution | 1-km resolution | 0.25-degree resolution | 4-km resolution |
| Ground-Truthing Correction | No ground-truthing adjustments | Ship and buoy ground-truthing | Jan 2002 – Dec 2010 data record (108 months) |
| Time Period | Feb 2000 – Dec 2009 data period (119 months) | Feb 2000 – Dec. 2010 data record (131 months) | Rain gage ground-truthing bias adjustment |
| Data Source | http://mrtweb.cr.usgs.gov/ | http://podaac.jpl.nasa.gov/ | University of Okla |

### 3.3.1 Enhanced Vegetation Index (EVI)

EVI data were obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument onboard the TERRA satellite (Table 3-1). Launched in December 1999, MODIS has been capturing high-quality multispectral data used to estimate vegetation density. The MODIS EVI product provides much higher spatial resolution than the Advanced Very High Resolution Radiometer (AVHRR) Normalized Difference Vegetation Index (NDVI) product, while matching its daily temporal resolution. The MODIS EVI calculates vegetation density in similar ways as NDVI. However, it corrects for a number of atmospheric particle distortions and saturation effect, has a greater optical penetration into canopies, and thus offers more robust reconstruction than NDVI in viewing areas with high biomass (like forests) (Chuvieco and Huete,

2010). In this study, monthly 1 km resolution EVI data were obtained from the MODIS Online Reprojection Tool for the period February 2000 through December 2009.

### 3.3.2 Precipitation

Level 4 NEXRAD data, processed by the University of Oklahoma, yielded precipitation estimates from the ground-based NEXRAD radar coverage with bias-adjustments from ground-truthing rain gauge sites. The original precipitation data in 4 km resolution were downscaled to 1 km resolution using cubic interpolation techniques and gridded over each site to match the scale and spatial extents of the EVI data. Data were available from February 2002 through December 2010.

### 3.3.3 Sea Surface Temperature

SST data were obtained from the Group for High Resolution Sea Surface Temperature (GHRSST) global Level 4 SST analysis. This product produces a daily 0.25-degree grid of SST at the NOAA National Climatic Data Center and uses optimum interpolation (OI) data from the 4 km AVHRR Pathfinder Version 5 time series. The data are bias-adjusted using in-situ ship and buoy observations. The resolution was resampled at a 0.5-degree scale using cubic interpolation to obtain more computationally efficient analysis of the Ocean boundaries. Data from 2000 through 2010 were used in the analysis.

## 3.4 Modeling Scheme

Identifying climate teleconnections conducted in this study expands upon the work performed by Huber and Fensholt (2011), who demonstrated strong teleconnection signals

between the African Sahel region and portions of the Atlantic and Pacific oceans; however, previous studies were limited to linear relationships between SST and greenness, discounting the nonlinear patterns that may influence teleconnection signals (Huber and Fensholt, 2011). In the current study, a linear correlation method was used as a simple initial screening process to identify indexed regions with strong sea-land correlation in Atlantic and Pacific oceans. Regions of these oceans with strong correlation for a period of at least 3 months are then extracted as individual SST indices followed by a stepwise regression analysis (SRA). All extracted indices are then modeled as independent variables in the SRA against the greenness and precipitation dependent variables for each site. Indices without statistically significant predictive terms were discarded. The final SST indices were then characterized using wavelet spectral analysis to search for nonlinear and non-stationary signals of climate teleconnections associated with the three terrestrial sites. The benefit of using wavelet analysis is that the time-frequency relationship of multiple time series can be simultaneously analyzed in comparison to the linear correlation analysis that can only address the "partially explained" results in a correlation analysis.

### 3.4.1 Linear Correlation Analysis and Data Preprocessing

Pixel-wise linear correlation maps were developed between each terrestrial dataset and the Atlantic and Pacific SST anomalies using Pearson's coefficient of correlation:

$$r = \left( \frac{\sum_i^n \left[ \left( x_i - \mu(x) \right) * \left( y_{i+d} - \mu(y) \right) \right]}{\sqrt{\sum_i^n \left( x_i - \mu(x) \right)^2} \times \sqrt{\sum_i^n \left( y_{i-d} - \mu(y) \right)^2}} \right) \quad (1)$$

where $r$ is the correlation coefficient, $n$ is the number of observations in a time series, $x$ is a measurement of terrestrial dataset time series, $y$ is an SST measurement in time series, $\mu$ is the mean and $d$ is the time lag. Before calculating the correlation coefficients, all datasets were converted to monthly averages (SST and EVI) or monthly totals of precipitation. Both terrestrial and oceanic time series were converted to anomaly values to remove seasonality. Anomaly values were computed by subtracting the monthly median value from each pixel for each month over each raster image. Median values, instead of mean values, were used for calculating climatology values because the time series is less than the 30-year period, and therefore the mean value would give undue weight to outliers over a shorter period as indicated by the World Meteorological Organization (Huber and Fensholt, 2011). For further noise reduction, all time series were smoothed using a 3-month moving-average filter. Each terrestrial dataset was reduced to a single time series by calculating the area-averaged value (Figures 3-4 and 3-5).

Pixel-wise correlation maps for each time series dataset were computed for all lags 0–12 months. Regions of highly correlated areas were extracted as individual SST indices and then further screened using the SRA method.

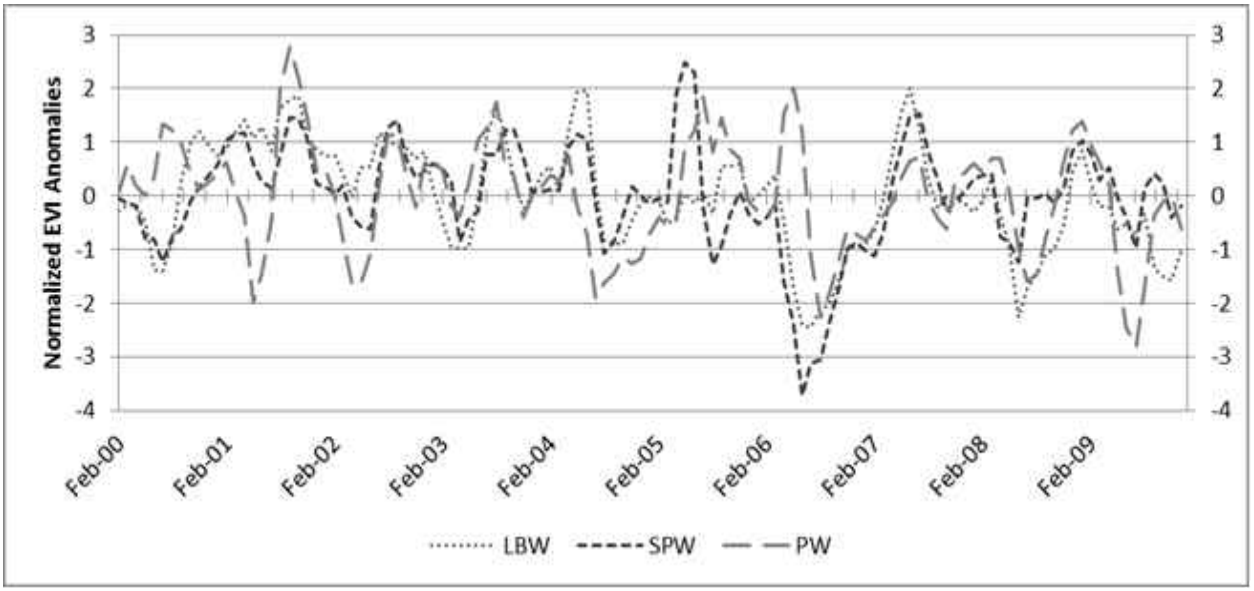


Figure 3-4: Normalized EVI anomalies in which greenness anomaly time series is less consistent, particularly the more coastal PW site, and shows a general decreasing trend.

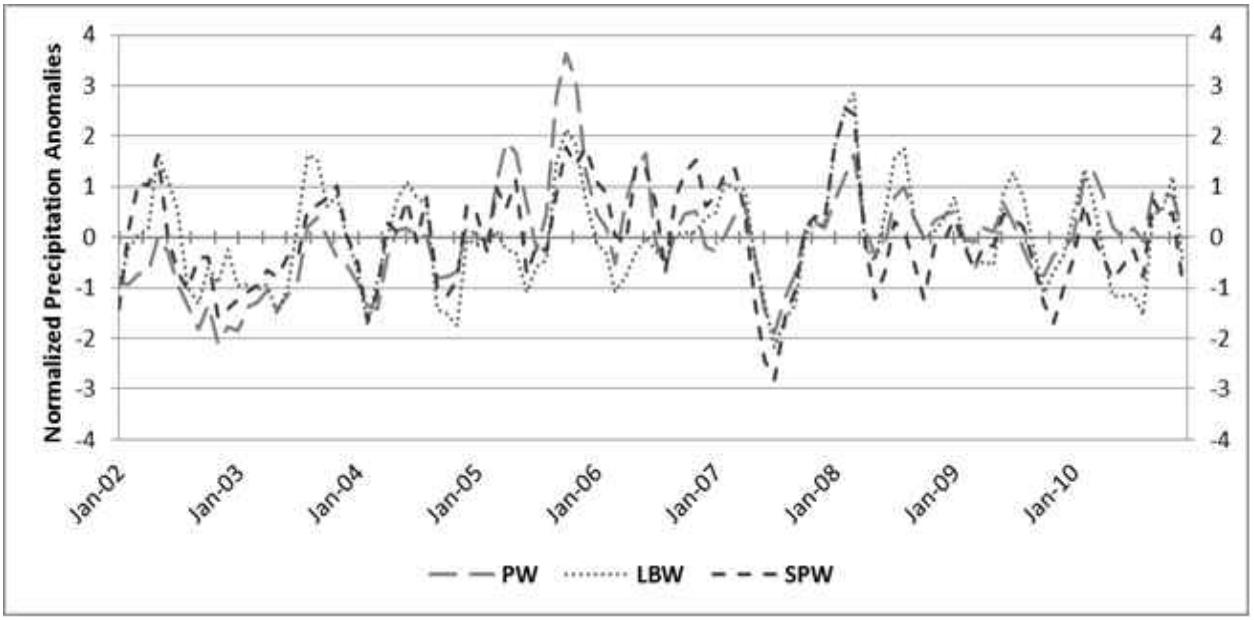


Figure 3-5: Precipitation anomaly in which the time series precipitation maps show a consistent pattern across all datasets, with a slow increasing trend across the time horizon.

#### 3.4.1.1 Stepwise Regression Analysis

SRA is a multilinear regression model that performs a series of regressions on an array of potential explanatory variables that are added or removed from the model at each step based on their statistical significance. Each step calculates the predictive value of a variable's coefficient estimate from its F-statistic and tests against the model with and without that potential parameter. The method is concluded when no additional variables can improve the model. Each extracted SST index was included as a potential explanatory variable in predicting greenness and precipitation at each site, whereas indices not found to be significant were discarded. A variable lagged SRA analysis was performed to find the best model with the highest $R^2$ value, which is considered the most significant time lag between the SST indices (independent variable) and terrestrial dataset (dependent variable), leading to the generation of the best causal relationship for explaining the teleconnection signal propagation with respect to multiple index regions in both Pacific and Atlantic oceans.

### 3.4.2 Wavelet Analysis

Wavelet analysis decomposes discrete time series observations into the time-frequency domain. The spectral analysis allows the determination of dominant localized variations of power (i.e., where the variance of the time series is largest for a given frequency) (Keener et al., 2010). By decomposing a time series into the time-frequency domain, one can determine both the dominant modes of variability and how those modes vary in time (Torrence and Compo, 1995). In this study, wavelet analysis was used to quantify and visualize statistically significant

changes in monthly North Atlantic and Pacific SST transients and the terrestrial greenness as well as precipitation over the three pristine forested sites.

3.4.2.1 Continuous Wavelet Transform

Assume a time series, $x_n$, with equal time spacing $\delta t$ and $n = 0 \ldots . N - 1$, and a wavelet function, $\psi(\eta)$, that depends on a nondimensional time parameter $\eta$. To be a legitimate wavelet, the function must have a zero mean and be localized in both time and frequency space (Farge, 1992). The Morlet wavelet used in this study is a wavelet consisting of a plane wave modulated by a Gaussian:

$$\psi(\eta) = \pi^{-\frac{1}{4}} e^{i\omega_o \eta} e^{-\frac{\eta^2}{2}}, \quad (2)$$

where $\omega_0$ is the nondimensional frequency.

Morlet wavelets are non-orthogonal, complex functions that can be used with the continuous wavelet transform $W_n(s)$. The continuous wavelet transform (CWT) of a discrete sequence $x_n$ is the convolution of $x_n$ with a scaled and translated version of $\psi(\eta)$, and is given by:

$$W_n(s) = \sum_{n'=0}^{N-1} x_{n'} \psi^*[(n' - n)\delta t / s] \ , \quad (3)$$

where * is the complex conjugate, *s* is the wavelet scale, $x_n$ is the discrete sequence, *s* is the scale, *n* is the localized time index, $n'$ is the translated time index and $\psi$ is the normalized wavelet. To approximate the CWT, the convolution should be performed N times for each scale, where N is the number of points in the time series. By choosing N points, the

convolution theorem allows all N convolutions to be performed simultaneously in Fourier space using a discrete Fourier transform. The discrete Fourier transform of $x_n$ is:

$$\hat{x}_k = \left(\frac{1}{N}\right) \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i k n}{N}}, \quad (4)$$

where k=0…N-1 is the frequency index. In the continuous limit, the Fourier transform of a function $\psi\left(\frac{t}{s}\right)$ is given by $\hat{\psi}(s\omega)$. By the convolution theorem, the CWT is the inverse Fourier transform of the product:

$$W_n(s) = \sum_{k=0}^{N-1} \hat{x}_k \hat{\psi} * (s\omega_k) e^{i\omega_k n \delta t}. \quad (5)$$

The wavelet power spectrum is defined as $|W_n(s)|^2$ and the amplitude at each point, $|W_n(s)|$, can be found. The univariate wavelet analysis was performed using the MATLAB WAVETEST script provided by Torrence and Compo (1995).

### 3.4.2.2 Cross-Wavelet Analysis

To demonstrate the teleconnection signal between two time series, a cross wavelet transform (XWT) is performed. A thorough description of cross wavelet analysis is provided by Torrence and Compo (1995), with practical applications and MATLAB scripts provided by Grinsted et al. (2004).

Given two time series *X* and *Y* with wavelet transforms $W_n^X(s)$ and $W_n^Y(s)$, the XWT defined as $W_n^{XY}(s)$ is:

$$W_n^{XY}(s) = W_n^X(s) W_n^{(Y*)}(s), \quad (6)$$

where * is the complex conjugate. The XWT identifies regions in time frequency space where the two time series show high common power and significance. Two time series with a significant XWT signal and relationship can suggest causation and thus potential teleconnection.

#### 3.4.2.3 Wavelet Coherency Analysis

The wavelet coherency transform (WTC) identifies regions where multiple time series covary in frequency space, but they do not necessarily have a high common power. WTC therefore demonstrates local covariance and will typically have more significant areas than XWT. The wavelet coherency is defined by Grinsted et al. (2004):

$$R_n^2(s) = \left( \frac{\left| S\left(s^{-1} W_n^{XY}(s)\right) \right|^2}{S\left(s^{-1}|W_n^X(s)|^2\right) \cdot S\left(s^{-1}|W_n^Y(s)|^2\right)} \right), \quad \textbf{(7)}$$

where S is a smoothing operator. This equation is notably similar to that of traditional correlation coefficient, but it is localized in time-frequency space.

Similarly, an XWT and WTC can be made between two time series datasets. The XWT may enable the identification of high common power between two separate time series and can suggest possible teleconnections between multiple regimes. The WTC may show how each time series covary, but the regions do not necessarily have high common power.

## 3.5 Results and Discussion

### 3.5.1 Linear Correlation and Wavelet Analysis of Anomaly Data

Across all three study sites, the greenness anomaly time series (Figure 3-4) is less consistent than precipitation, particularly in the more coastal PW site, and shows a general

decreasing trend. The time series precipitation maps (Figure 3-5) show a more consistent pattern across all datasets, with a slow increasing trend across the time horizon. The time series of greenness versus precipitation over each site (not shown) display little consistency, with periods of low greenness associated with high precipitation and vice versa. A basic visual interpretation of the graph comparison is generally unhelpful, particularly if underlying low frequency relationships are to be understood. Wavelet analysis permits the underlying modes of variability between the time series to be observed.

To demonstrate wavelet analysis, two CWT expanded time series for SPW greenness and the Index 1 SST (east coast of U.S./Canada as shown in figure 3-9) were carried out (see Figure 3-6). Both time series include the raw, unfiltered time series record, averaged spatially over their respective regions. The SPW greenness graph shows a strong signal (dark red) centered on the 12-month frequency, which is consistent with the annual cycle of vegetation growth. The graph also shows a moderately strong signal (green and yellow region) centered on the 6-month frequency, a consistent finding considering that the vegetation density in the autumn and spring can appear similar, whereas summer and winter greenness appear different, causing frequency strengths roughly half of those expected at the 12-month period. Notice, however, that there is no significance line (bold black line) around the 6-month frequency. The "U" shaped region of the graph is consistent from one image to the other and is determined based on the length of the time series. In order for a wavelet to determine the power of each period, the non-dimensional wavelet expands to encompass larger periods and contracts to encompass smaller periods. Thus, larger periods suffer from "edge effects" where the larger

wavelets have a much smaller window over which they can completely propagate through the time series; therefore, zeroes are “padded” to either end of the time series to compensate. These areas of the graph, indicated by the opaque regions and known as the Cone of Influence (COI), are not fully reliable because the wavelet is influenced by the artificial zero values. Longer periods have a very narrow region of the graph where the data are considered valid, due to the edge effects. SST (bottom image) has a similar appearance to greenness, which is consistent with the annual warming and cooling of the oceans.

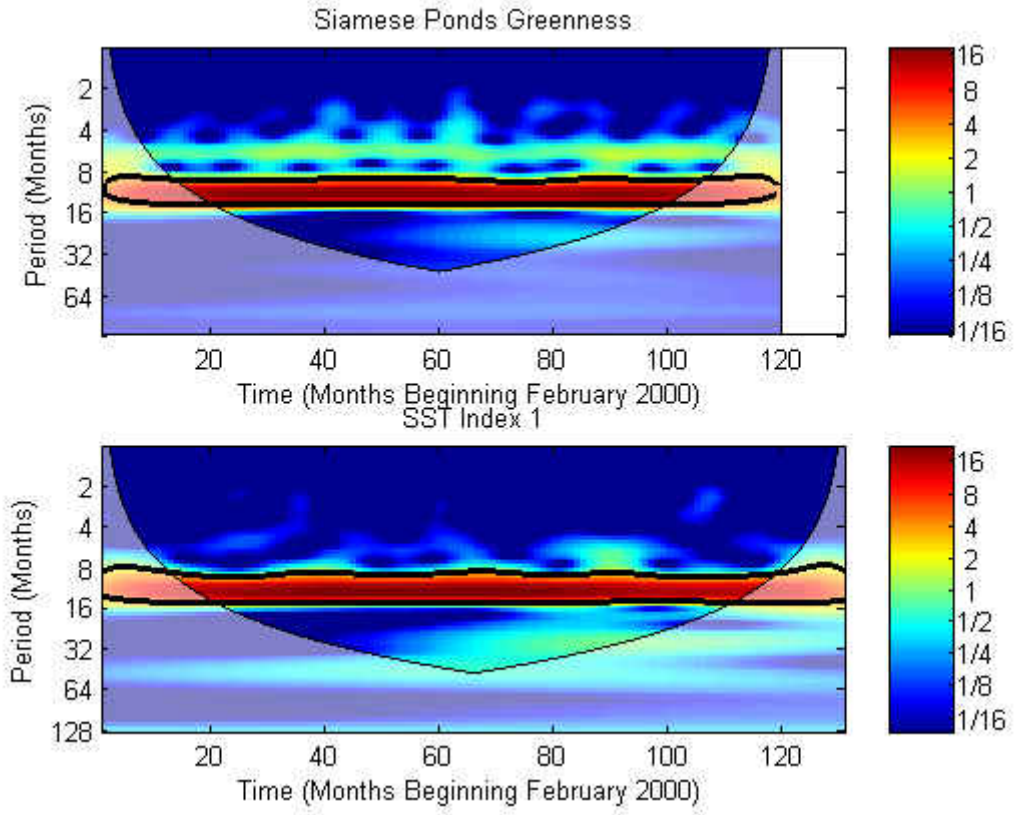


Figure 3-6: CWT of Siamese Ponds greenness (top image) and SST Index 1 (bottom image) datasets. For both images, the x-axis is the time (in months) beginning February 2000 and the y-axis is the frequency period (in months). The opaque regions are areas of the graph subject to edge effects of wavelet analysis. The color indicates wavelet power, with high power appearing dark red. The bold black line indicates regions of statistical significance.

Lagged correlation anomaly maps from SPW greenness and PW precipitation show strong correlation exists in large regions throughout the northern Atlantic and Pacific oceans (Figures 3-7 and 3-8). Time lag of 0–12 months refers to time of SST leading both precipitation and greenness. Across all correlation maps, minimum spatial correlation across ocean boundaries ranges from 12–45%, exceeding the 10% correlation anticipated with a two-tailed 5% probability threshold, indicating that the correlation patterns do not reflect random variability. A general pattern in the anomaly maps indicates a mixed positive and negative correlation throughout both oceans from lags 0–7 months; specifically a strong negative (positive) correlation dominates greenness (precipitation) during the longer lag periods. For LBW (not shown) and SPW, greenness shows stronger correlation trends than precipitation, as determined by the extent of significant correlation regions. This trend is reversed for PW, which shows a strong precipitation correlation, particularly in the North Atlantic.

The analysis results further show particular areas of the Atlantic and Pacific oceans are more significant in correlation than other areas. In the Atlantic, regions south of Greenland show strong, long-term evidence of correlation consistent with the NAO teleconnection regions (Barnston and Livezey, 1987). The greenness and precipitation correlation regions in the Atlantic at the 12-month time lag reveal a distinct layered tripole of correlation consistent with the NAO meriodonal variability reported in IPCC 2007, with a strong negative correlation around Greenland and north of the 50°N latitude, followed by a strong positive region of correlation between 30°N–40°N before transitioning to negative again at 25°N. This 12-month tripole pattern is aligned with the 1-year memory of the winter NAO,

where observations have shown that 25% of the winter NAO variance can be predicted from the preceding seasonal SST pattern (Czaja and Frankignoul, 1999). This evidence is consistent with an oceanic feedback to atmospheric forcings, although it is not clear what effect this memory has on greenness and precipitation responses.

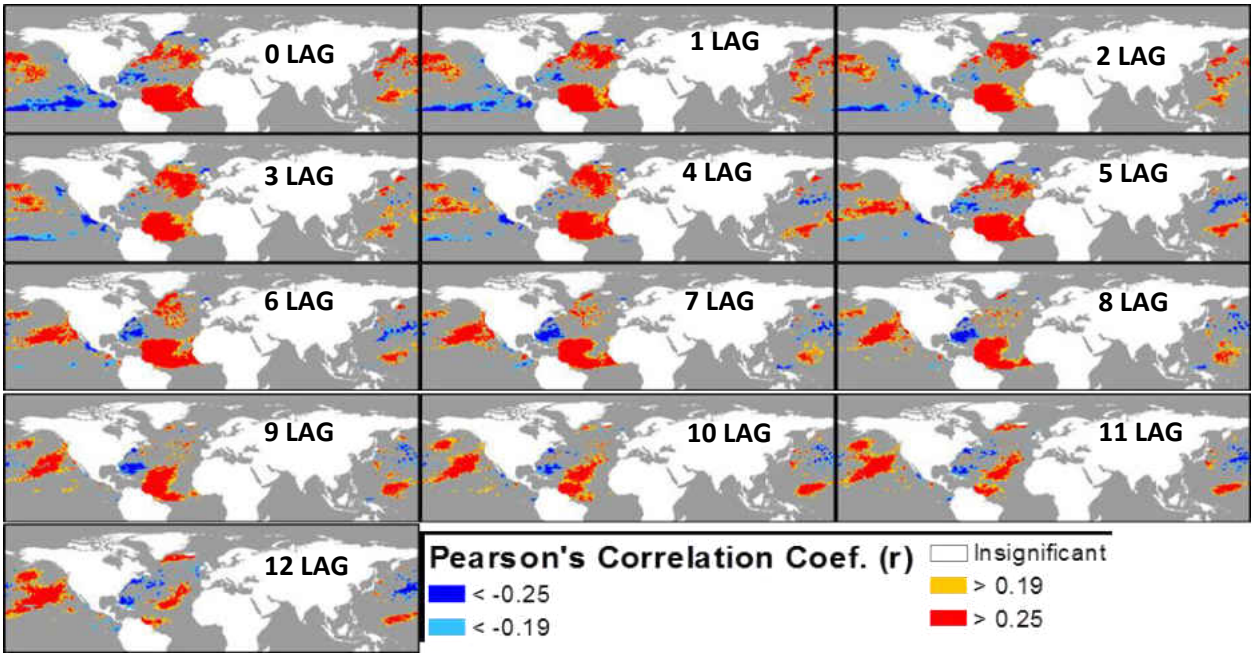


Figure 3-7: PW precipitation and SST correlation maps for all lags 0-12 months. For all images, dark red represents statistically significant regions of correlation at the 99% confidence level, dark blue represents anticorrelation at the 99% confidence level with yellow and light blue representing significant positive and negative correlation at the 5% confidence level, respectively.

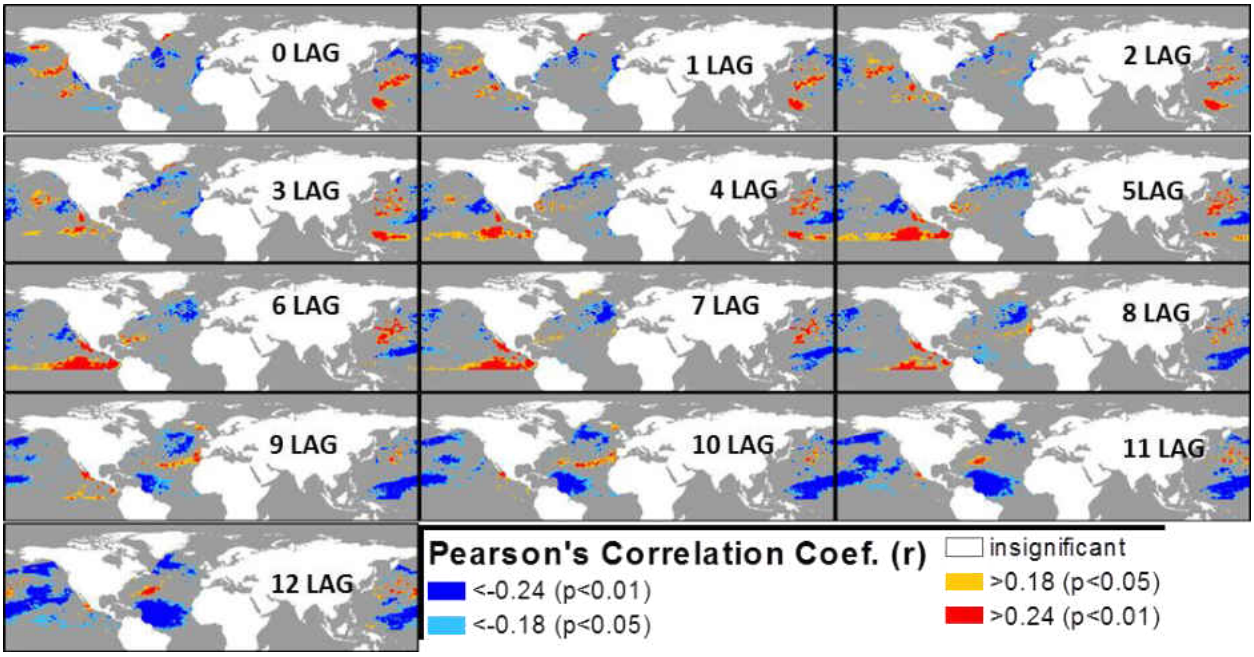


Figure 3-8: SPW EVI and SST correlation maps for all lags 0–12 months. For all images, dark red represents statistically significant regions of correlation at the 99% confidence level, dark blue represents anticorrelation at the 99% confidence level with yellow and light blue representing significant positive and negative correlation at the 5% confidence level, respectively.

In the Pacific Ocean, the equatorial regions of the ocean display the broadest extent of correlation, generally positive during the middle time lags and negative during later time lags. This strong correlation in the equatorial Pacific Ocean is consistent with the known ENSO climatic impacts in the contiguous U.S. Strong correlation also is observed in the more northern latitudes of the Pacific in regions with known PNA teleconnections. In the lag periods of 10–12 months, greenness maps show similar strong anticorrelation with SST anomaly in the latitudes of 0–30°N, slight positive correlations from 30–45°N and strong anticorrelation again from 45–60°N.

The areas of high correlation in both oceanic regions seem to follow particular time-space patterns and slowly "drift" from one time lag to the next. These areas are noticeable in the mid-latitudes of the Atlantic Ocean and are consistent with the eastward advection of anomalous heat patterns driven by the Gulf Stream, which repeat over a 6-year period (Marshall et al., 2001). A detailed evaluation of the relationship between oceanic currents and teleconnection signals is beyond the scope of this study, however.

From the correlation maps, regions with long-term (more than 3 months) and consistently high correlation were extracted from the SST anomaly dataset (Figure 3-9) and further converted to single indices using an area-weighted average. Index 1, located in a teleconnection region known as the North Atlantic Ridge (RDG) identified by (Cassou et al., 2004), is a region closest to the sites within a consistent negative greenness correlation in early time lags and a strong positive precipitation correlation. Indices 2–4 are regions south of Greenland that coincide with the known NAO SST anomalies. Index 3 displays a strong positive

correlation with greenness in the early months in SPW and LBW, and indices 2 and 4 show strong negative correlations with greenness in middle-late time lags with SPW and LBW and strong positive correlation with precipitation specifically with the PW site. In the Pacific, Indices 5, 7 and 8 represent SST anomalies located along the equatorial line within the limits of the known ENSO Niño Indices 3, 3.4 and 4, respectively. These indices all demonstrate strong positive correlation with greenness, with Index 8 correlating strongly at lags 0–5 months and Indices 5 and 7 at slightly later time lags. Index 6 is not directly associated with traditional teleconnection regions. Just off the western coast of Mexico, Index 6 exhibits one of the most consistently positive correlations with greenness for SPW from time lags 1–10 months and from 3–9 months for LBW (not shown). The area northeast of Brazil was not included in the index extractions due to the generally long lag period needed to establish a correlation.

The extracted eight indices were screened using a lagged SRA model to determine the degree at which an index, or combined indices, could explain the most vegetation and precipitation dynamics at each site. The best model demonstrating the highest $R^2$ value for each variable (greenness and precipitation) reflects the number of months that greenness and precipitation lag SST (Table 3-2). Only those indices that were included in the SRA model provided lag information. The reported $R^2$ values from the SRA models show values generally consistent from 0.30–0.50, with a PW greenness outlier of 0.07. Similar studies, such as Huber and Fensholt (2011), demonstrate multiple ocean influences of >0.50. These studies, however, typically use the favorable winter-only period to calculate correlation because this season has been demonstrated to show the most consistent interannual variability. The remaining seasons

can be severely muted by the seasonal noise in the data. An $R^2$ value of 0.50 across a continuous monthly time series shows the benefit of site screening and selection of pristine areas.

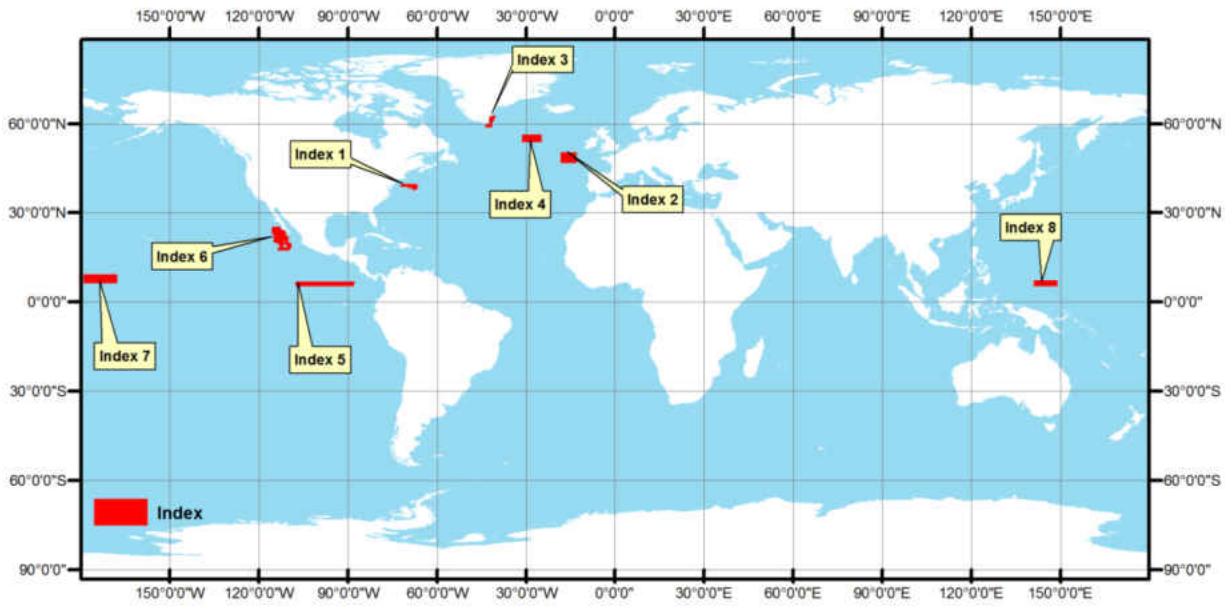


Figure 3-9: Ocean indexed sites identified with high correlation for teleconnection signal propagation.

Table 3-2: Stepwise regression analysis with reported indexed regions' t-statistic. Indexed regions with a t-statistic of a magnitude of 2 or higher, corresponding to a probability of <5%, were included in the final model, with the $R^2$ value reported in the final column. The numbers in parenthesis correspond to the number of months in which greenness or precipitation lag SST.

| Site / Index | Index 1 | Index 2 | Index 3 | Index 4 | Index 5 | Index 6 | Index 7 | Index 8 | r | $r^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| SPW Greenness | **-3.4 (2)** | **-3.8 (7)** | **2.8 (0)** | -1.3 | -1.2 | 0.7 | -0.5 | **5.5 (3)** | 0.71 | 0.50 |
| SPW Precip | **3.5 (2)** | **3.4 (2)** | -0.8 | **2.6 (4)** | -1.5 | 0.3 | **4.2 (8)** | 0.4 | 0.65 | 0.42 |
| LBW Greenness | **-2.3 (3)** | -0.9 | 0.9 | -0.4 | **-5.8 (0)** | **5.4 (6)** | 0.2 | 1.2 | 0.63 | 0.40 |
| LBW Precip | 1.6 | **3.0 (2)** | -0.4 | 1.9 | **2.2 (3)** | -0.4 | **-4.8 (9)** | 1.8 | 0.56 | 0.31 |
| PW Greenness | **-2 (3)** | 0.96 | 0.65 | 0.38 | **-3.02 (0)** | -1.47 | -0.86 | 1.97 | 0.26 | 0.07 |
| PW Precip | **4.9 (2)** | **2.6 (2)** | 0.2 | 1.8 | -1.0 | 1.2 | 1.4 | 1.6 | 0.64 | 0.41 |

All indices were used in at least one SRA model (Table 3-2); however, the Atlantic Ocean indices have greater influence on the three sites in the northeast U.S., as determined by the number of included indices in all SRA models. Indices 1 and 2, located off the U.S. northeast coast and the Azores, respectively, show the most influential ocean region studied across all sites. Index 1 is used in all greenness models, with a consistent 2-month lag period, while Index 2 is used in all precipitation models with a 2-month lag period. These findings agree with earlier studies that the New England climate is linked with NAO and regional New England SST (AIRMAP, 2004; Marshall et al., 2001). The 7-month lag for SPW greenness and Index 2 indicate that precipitation explains a low amount of the greenness variability, a finding confirmed by previous studies between northern U.S. forested greenness and precipitation (Zhou, 2003),

where temperature patterns are the primary drivers of greenness. Index 5, associated with the Niño 3 region, shows strong correlation and a 0 lag with LBW and PW greenness. A 0-month, "instant" correlation between the northeast U.S. and the eastern equatorial Pacific suggests that atmospheric forcings cause the coincident anomalous behavior. Such behavior is possible during periods of strong ENSO where upwelling of cold waters in the eastern Pacific are driven by strong trade winds, which are also known to affect mid-latitude westerlies and the location of the jet stream over the northeast U.S. The short lags of Index 5 greenness are juxtaposed with the 8–9 month lags of Index 7 precipitation for both the SPW and LBW sites. These time lags in the Niño 3.4 region reflect a longer, more remote relationship with storm track patterns across the northeast U.S.

For greenness, the SPW displayed the best model, with 50% of the variability being explained by the combined Indices 1, 2, 3 and 8 at time lags varying from 0–7 months. This model shows the complex relationship that the New England climate has with varying teleconnection regimes, and that multiple ocean basins play important roles in the local ecosystem dynamics. For precipitation, each site showed a fairly consistent SRA model, with $R^2$ values ranging from 0.31–0.42. Additionally, Indices 1, 2 and 7 demonstrate consistent skill across the study regions, suggesting that predictive models may be possible.

### 3.5.2 Wavelet Analysis Results

For all wavelet maps, dark red indicates a positive correlation with statistical significance above the 99% confidence level, and dark blue indicates negative correlation above the 99% confidence level. Light blue and yellow indicate negative and positive correlation above the 5%

confidence level, respectively. CWT images for each terrestrial and oceanic index (Figures 3-10 and 3-11) show the univariate time series transform of the anomalous datasets. These graphs expand the one-dimensional time series into a two-dimensional time-frequency image. The x-axis represents the date, in years, beginning in February 2000 and ending in December 2010. The y-axis is the period (frequency) in months. The color levels are indications of wavelet power, with red indicating high power and blue indicating low power. The opaque color region of the graph indicates the influence of edge effects in wavelet analysis, a result of padding each time series with zeroes. Bold black lines indicate areas of statistical significance at the 95% confidence level.

The CWT images reveal different patterns between each dataset. For terrestrial data, the greenness images show the most consistency between all sites and are characterized by weak high-frequency significance regions beginning around the 4-month period and increasing in power at lower frequencies of 12 months and lower. Each site shows a clear significance region at the 16–24 month period, beginning in the early months of 2005. For precipitation, the CWT images also show high-frequency significance from the 4–8 month period, with SPW and LBW displaying a strong 10–16 month period beginning in mid-2006 to 2008.

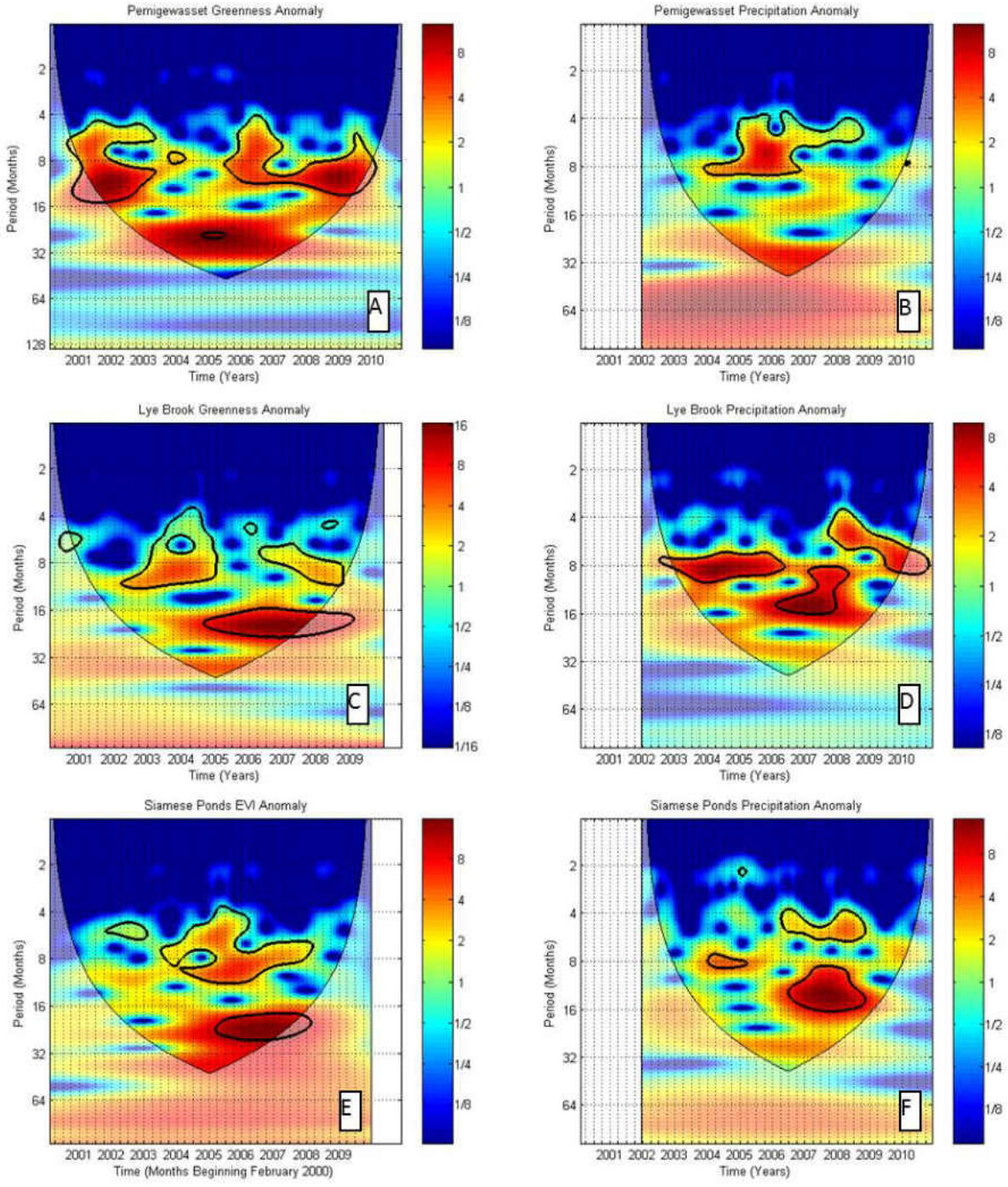


Figure 3-10: CWT images of the terrestrial index time series: (A) PW Greenness anomaly; (B) PW Precipitation anomaly; (C) LBW Greenness anomaly; (D) LBW Precipitation anomaly; (E) SPW Greenness anomaly; (F) SPW Precipitation anomaly.

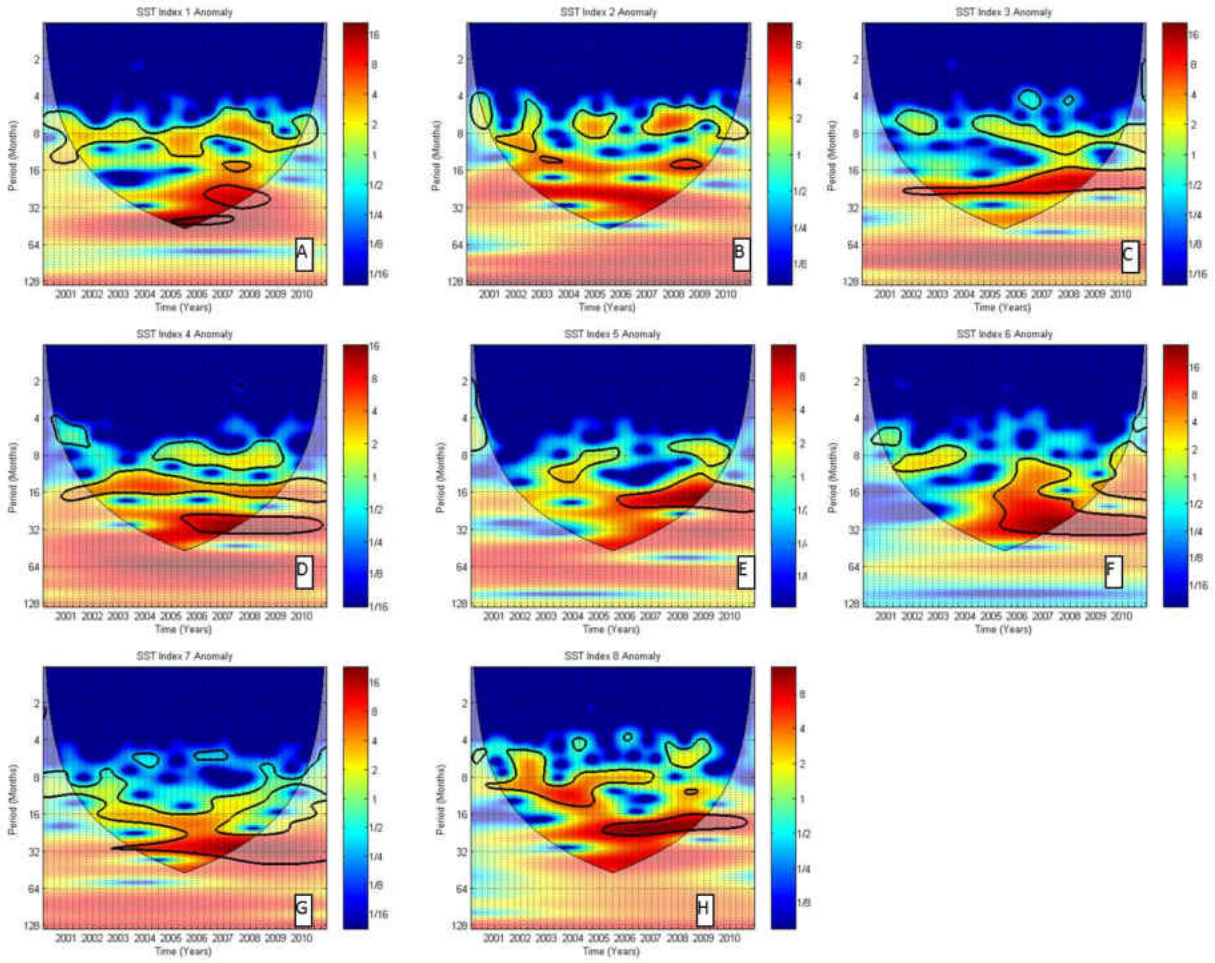


Figure 3-11: CWT images of the oceanic index time series: (A) Index 1 anomaly; (B) Index 2 anomaly; (C) Index 3 anomaly; (D) Index 4 anomaly; (E) Index 5 anomaly; (F) Index 6 anomaly; (G) Index 7 anomaly; (H) Index 8 anomaly.

The oceanic CWT images also show consistent signals throughout most of the extracted indices. Generally, a weak yet significant high-frequency signal is observed between the 4–8 month period followed by a much stronger significant signal between 16 and 32 months. As with the terrestrial CWT images, the low-period signal exists between mid-2005 and 2008; however, this pattern is primarily observed with the Pacific datasets. Interestingly, this time period is consistent with a change in the southern oscillation from an El Niño period to a

moderate La Niña one. A similar phase shift also exists during this period with NAO transitioning from a historically positive phase to a slight negative phase in 2005 followed by a more significant negative phase in 2008–2009. The general similarities between the terrestrial and oceanic CWT images, especially in the 2005–2008 low-frequency signals, suggest the existence of a low-frequency teleconnection.

Following the CWT anlaysis, the XWT and WTC images were analyzed to retrieve the most significant terrestrial–oceanic pairs identified in Table 3-2 and compared to the avaiable linear time series. WTC and time series graph of SP greenness and Index 3 are provided in Figure 3-12. The WTC image shows varying patterns of high-frequency coherency, specifically in the late 2003 and 2008 period, identified by regions A and B. The strong coherency can be seen on the time series graph, where the normalized anomalies closely match. The lower-frequency, 24-month period is consistent across the entire time domain, suggesting a biennial teleconnection response to greenness from the Index 3 SST. Similar patterns are also observed on the time series as the 24-month moving averages. Unlike the CWT, the WTC images display arrows to determine phase direction and lead–lag relationships. Arrows pointing to the right indicate that the two time series are in-phase; arrows pointing to the left indicate an anti-phase relationship; and arrows pointing up or down indicate a lead–lag relationship, with upward arrows indicating SST leading greenness or precipitation. The complete circle of the phase arrows equates to the period of frequency, meaning that phase direction along lower frequencies imply longer lead and lag times, with the exception of 0° and 180° reflecting absolute in-phase and anti-phase, respectively. The low-frequency signal in Figure 3-12

indicates that SST is in-phase and leads greenness, meaning that anomalously high (low) temperatures in Index 3 are associated with anomalously high (low) greenness patterns at SPW with a lead time of several months.

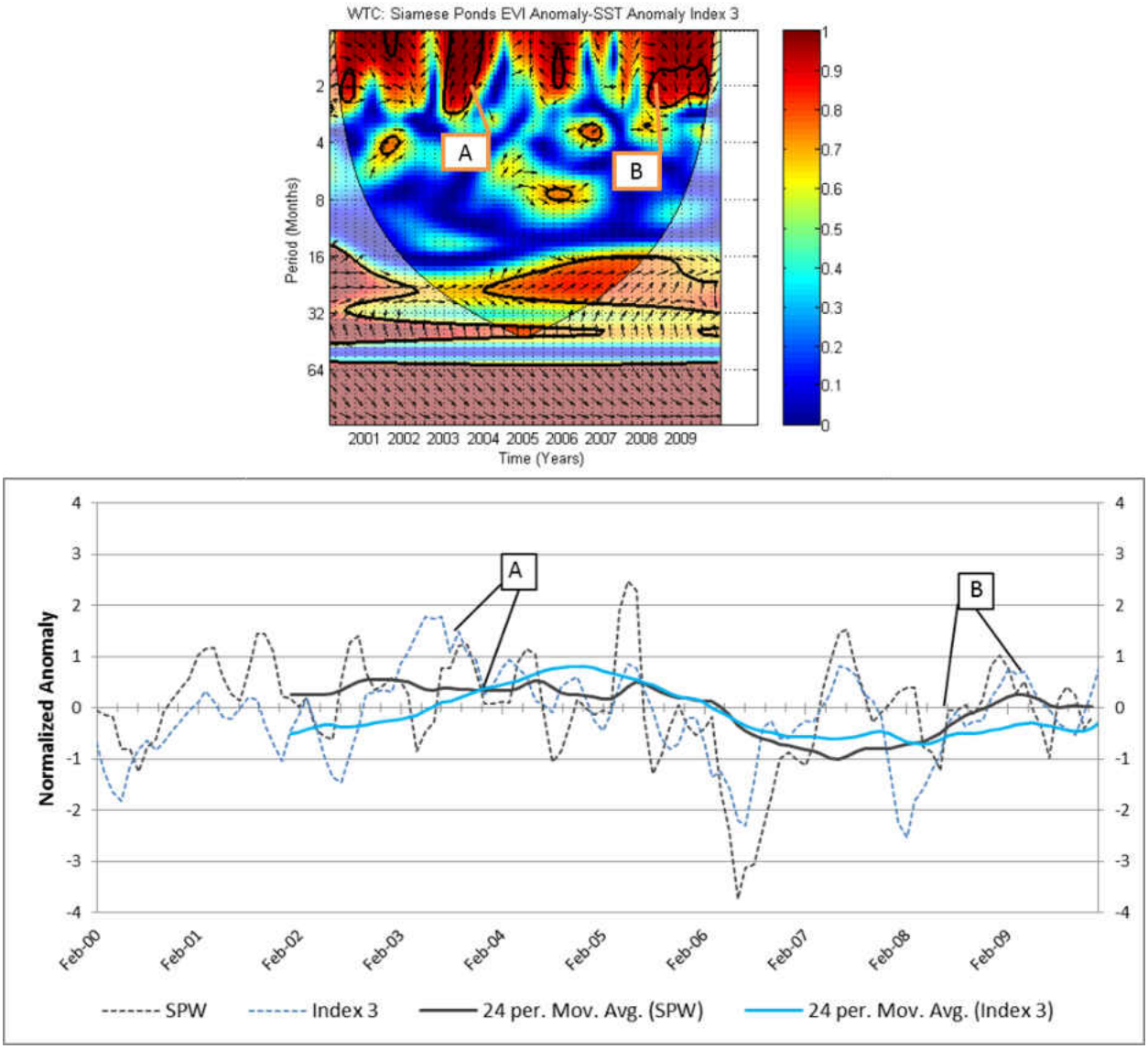


Figure 3-12: WTC image of SP greenness and SST Index 3 (top) and time series (bottom). Callouts A and B show two regions of high frequency coherency. Strong low frequency coherency is shown throughout the time domain in the 20–30-month period in the WTC image and in the 24-month moving average of the time series.

WTC image and time series of PW precipitation and SST Index 1 (Figure 3-13) show that, similar to SPW, periodic high-frequency coherency is present, but the more consistent period is found with lower-frequency bands, centered on the 30-month period. The time series moving average lines show a strong consistency with the WTC image.

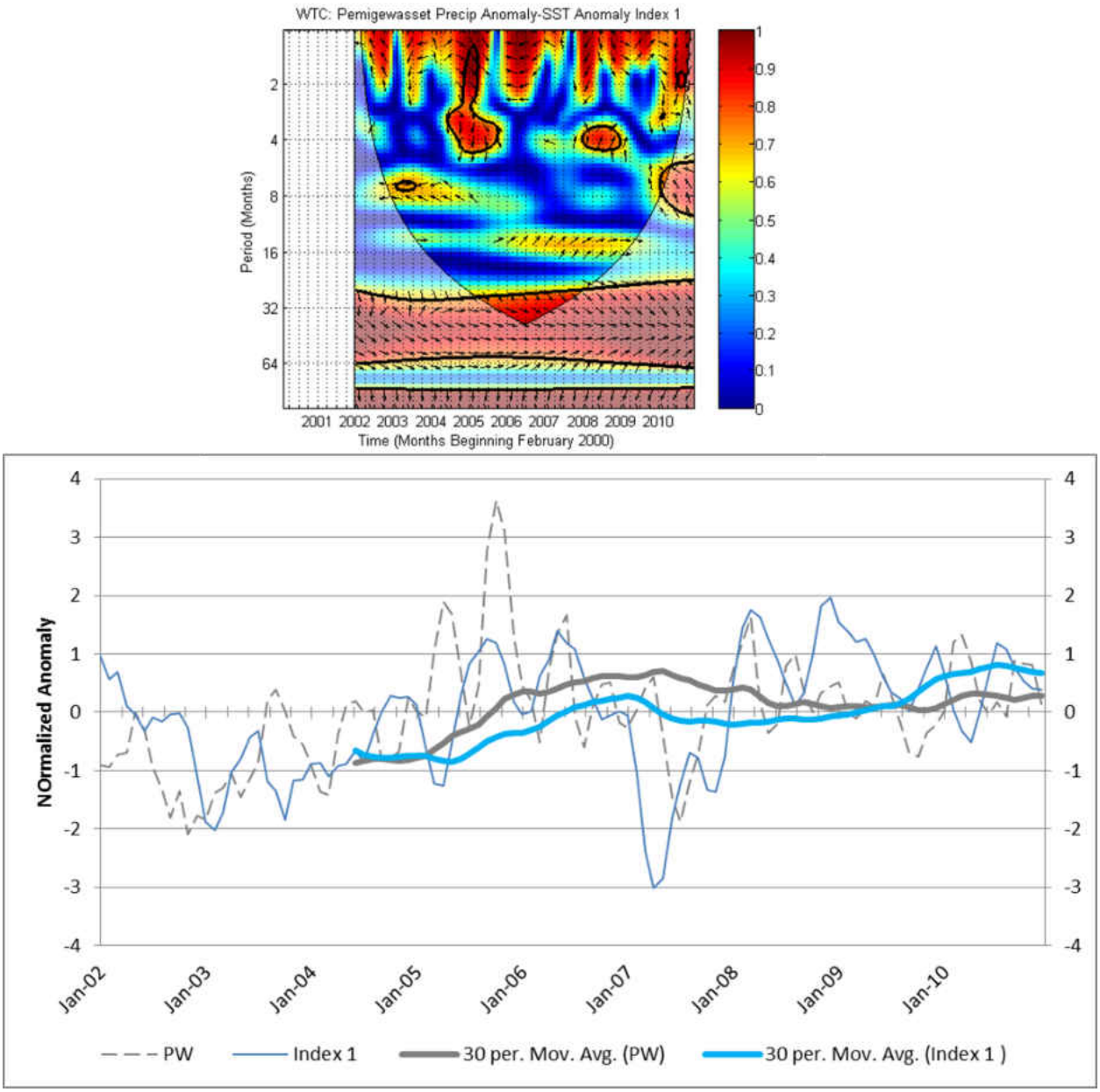


Figure 3-13: WTC image of PW precipitation and SST Index 1 (top) and time series (bottom). Strong low frequency coherency is shown throughout the time domain at the 30-month period in the WTC image and in the 30-month moving average of the time series.

The XWT and time series of PW greenness and the associtaed multivariate ENSO index (MEI) are shown with the time series displaying the Index 7 anomalies (Figure 3-14). The time series shows how the MEI and Index 7 (Niño 3.4 region) relate, with colder periods associated

with a negative MEI (La Niña) and warmer periods associated with a positive MEI (El Niño). The PW greenness shows an opposite trend in the time series, with El Niño periods associated with negative greenness trends and La Nina periods with high greenness. The time series further reveals that greenness patterns are particularly negative in the summer following a La Niña winter. Because La Niña periods are consistent with mild to warm winters, reduced summer greenness could be a response to early onset spring conditions or an increased water demand due to evapotranspiration rates. The XWT image confirms the negative phase trend of the time series, showing a distinctive anti-phase direction. Unlike WTC, the XWT shows common high wavelet power and not just local coherency. This image shows a strong interannual teleconnection across the entire time period, with significant strength beginning in 2004–2005 when the ENSO period began to shift toward a more dominantly La Niña period. The weak SRA model for PW greenness, however, suggests that predictive skill across this time domain is low.

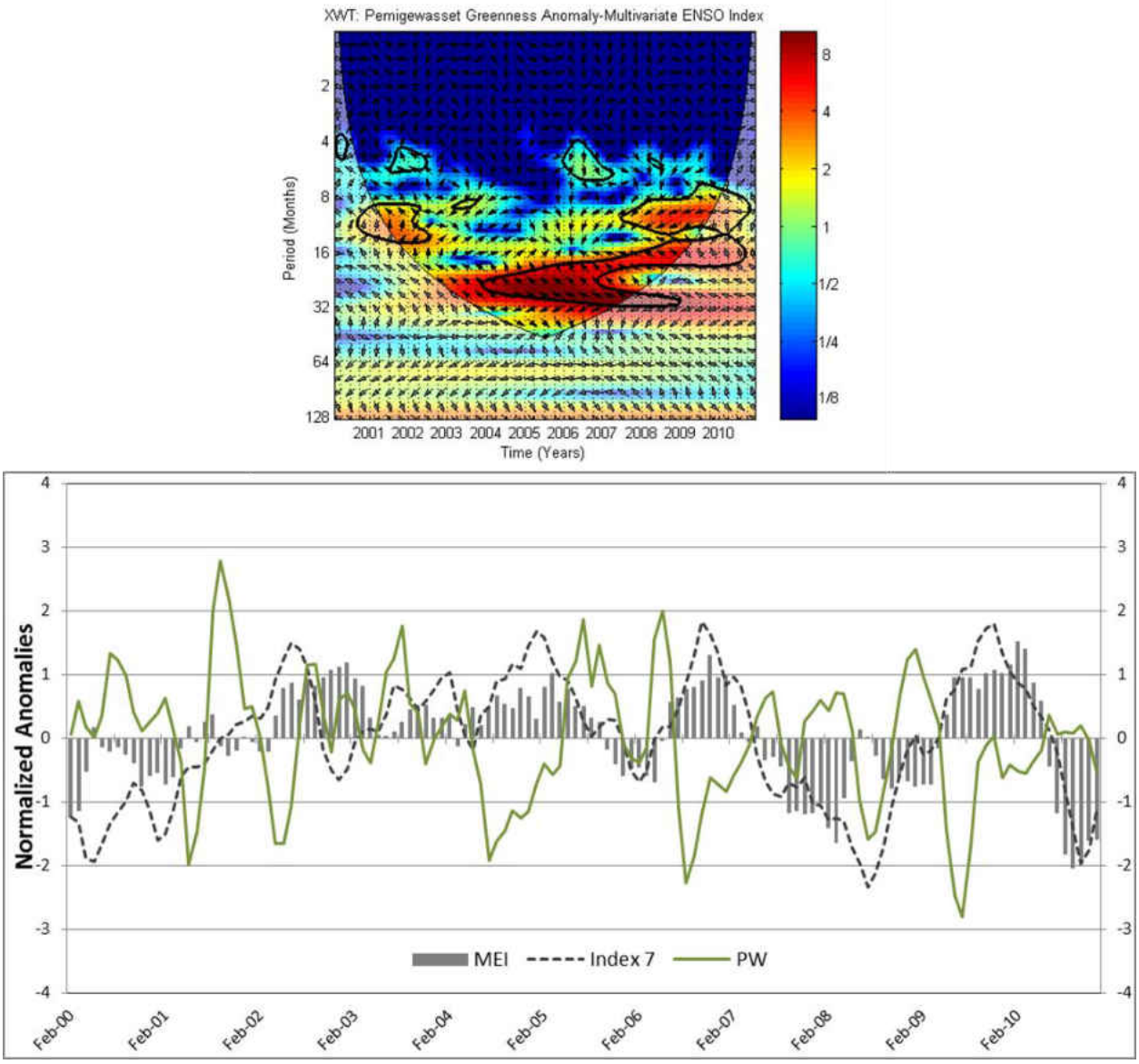


Figure 3-14: XWT image of PW greenness and MEI (top) and time series of PW greenness, MEI, SST Index 7 (bottom). Strong interannual wavelet power is shown throughout the time domain at the 20–30-month period in the XWT image with a consistent negative phase. Negative greenness summer events are common following a La Niña winter.

The LBW greenness and Index 6 WTC and time series shown in figure 3-15 (with SST shifted 6 months), show the closely related variability between regions of the Pacific and the northeastern climate patterns to be consistent with the Pacific-induced SST variability identified

by Enfield and Mayer (1997). This region is not known to be part of a tradional teleconnection region, and is located just north of the equator, off the west coast of Mexico. Further research with longer periods of data are needed to determine if this area shows teleconnection evidence across different terrestrial regions (Wunsch, 1999).

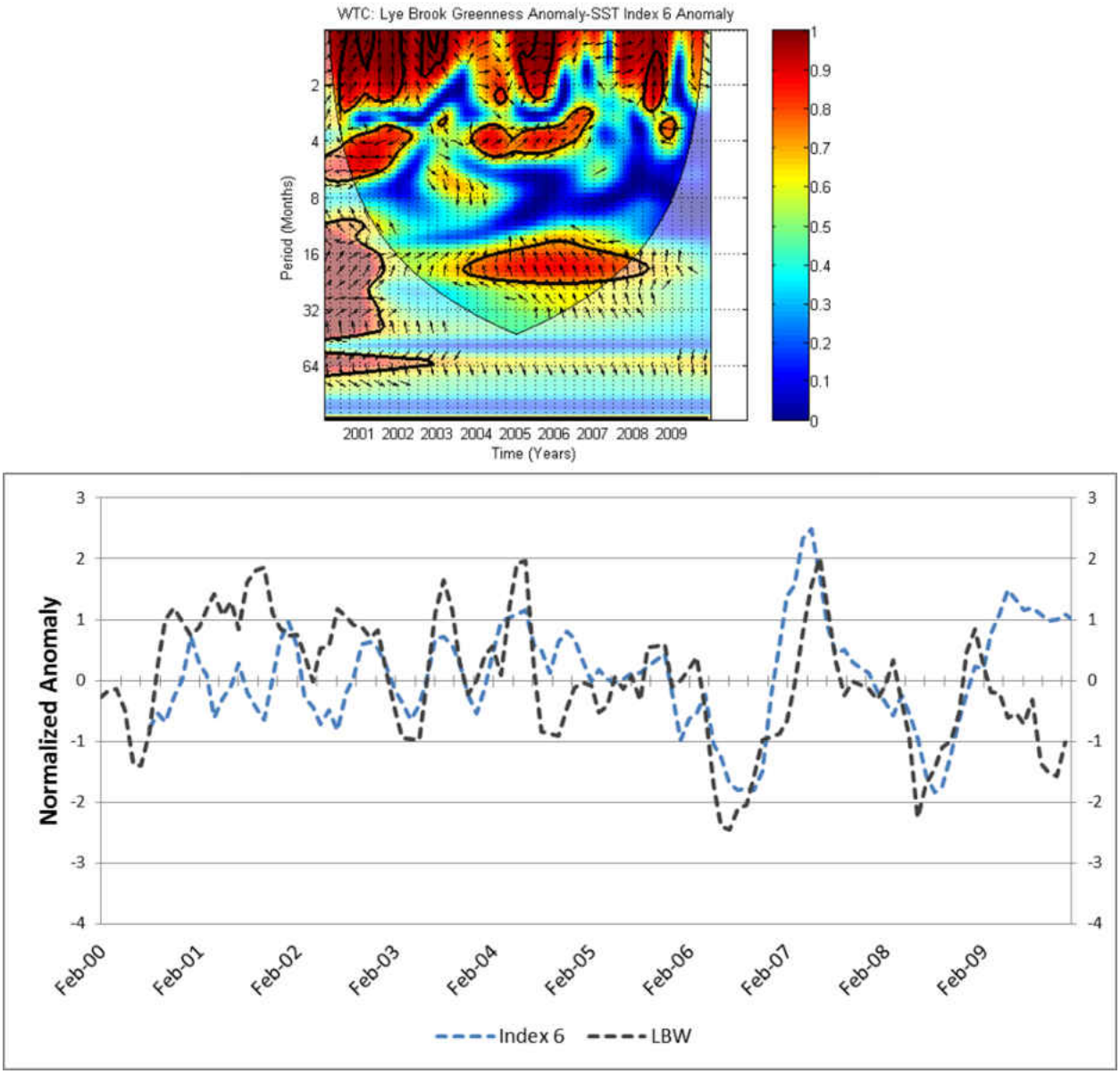


Figure 3-15: WTC of LB greenness and SST Index 6 anomalies (top) and time series with a 6-month shift for Index 6 (below).

Based on the sciences questions posed, both high- and low-frequency signals were identifeid between SST anomalies and greenness and precipitation. From the correlation maps, greenness and precipitation are shown to contain statistically significant relationships to both

known teleconnection regions (ENSO, NAO) as well as previously unknown teleconnection regions (i.e. Index 6). In the north Atlantic, higher-frequency seasonal time lags (1-3 months) appear to have the most significance on terrestrial greenness and precipitation, whereas in the north Pacific, lower-frequency time lags (6-9 months or longer) appear to be dominant. The wavelet images reveal that even lower-frequency periods, such as the biennial-triennial signal, may also play an important role in understanding the oceanic-based memory on terrestrial climate.

The precipitation patterns at the three sites were not able to clearly explain a significant porition of the greenness variability, suggesting that temperature may be the dominant physical process driving greenness trends. Finally, the Pacific and Atlantic-based teleconnections suggest that the combined ocean regions may aid in the predictability of local greenness and precipitation variability when accounting for multiple oceans and their associated memory effect on local climate conditions. In the following chapter, these underlying relationships will be further explored to develop teleconnection-based precipitation forecasting models.

# CHAPTER 4 : LONG-TERM TELECONNECTION IDENTIFICATION AND PRECIPITATION FORECASTING

In the previous Chapter, short-term climate teleconnection signals were identified between the northeastern U.S. and the northern Atlantic and Pacific basins. The findings suggest that nonlinear statistical relationships exist, which may be useful in climate prediction models. As is evident from the teleconnection maps, consistent regions in the northern Atlantic and Pacific basin associated with known teleconnection (ENSO, NAO, etc...) are present, and can explain a large portion of the greenness and precipitation variability within the three study sites. These dominant teleconnection regions are supplemented by regions of the oceans not typically associated with teleconnection influences. It is the intent of this chapter to establish both the leading and non-leading teleconnection patterns that are particular to the Adirondack region in the northeastern U.S. These teleconnections will be objectively identified through statistical relationships and compared against leading teleconnection regions, with the overall objective to produce a precipitation forecasting model based on teleconnection input.

To establish the occurrence of potential non-leading teleconnection patterns unique to a particular region, it is critical to understand the historical influence of leading teleconnection patterns on local climatic conditions. As discussed in Chapter 2, the northeastern U.S. has been well documented as having strong correlations between such teleconnection indices as the Pacific North American Pattern (PNA), ENSO and the North Atlantic Oscillation (NAO). The warm El Niño phase of the ENSO has been linked with cold, stormy winter weather in the

northeastern region and warm mild weather under La Niña conditions. The NAO, traditionally measured by sea level pressure differences between the Azores and Iceland, is the dominant pattern of Atlantic climate variability, and has been shown to have broad influence over Atlantic storm track, including nor'easters. PNA is a dominant extratropical teleconnection pattern in the Northern Hemisphere, characterized by above average SLP heights over the Hawaiian region and over the mountainous regions of western North America, as well as below average heights south of Alaska and over the southeastern U.S. The positive cycle of PNA has been linked to increases of precipitation over the northeastern U.S. by up to 15% over the last century (Beckage et al., 2008; Tang and Beckage, 2010). Figure 4-1 illustrates generalized locations of the leading teleconnection regions associated with ENSO, NAO and PNA (teleconnection regions over land masses are not shown).

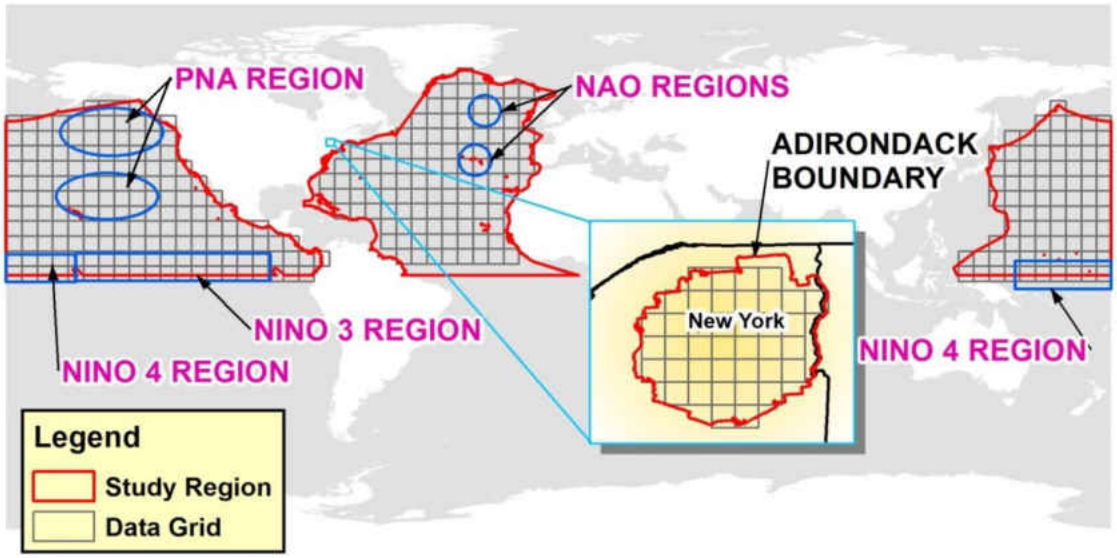


Figure 4-1: Gridded boundary regions for Adirondack State Park and the northern Atlantic and Pacific basins. Generalized leading teleconnection regions over oceanic areas are outlined in blue.

The following scientific questions are explored in the long-term investigation: (1) Do non-leading teleconnection signals exist between the Adirondack region precipitation and SST in the northern Atlantic and Pacific basins? (2) Can an ANN monthly precipitation prediction model be developed using only historic precipitation time series as well as teleconnection patterns in the northern Atlantic and Pacific oceans? (3) Does the ANN precipitation prediction model perform better than a statistical downscaling model for the same study region and time domain? This paper hypothesizes that both leading and non-leading teleconnection signals exist for the study region, and that a reasonable ANN long-term precipitation forecasting model can be developed using only historic precipitation and the objectively identified teleconnection inputs.

## 4.1 Data and Methods

### 4.1.1 Data

The precipitation dataset used for this study is the Climate Prediction Center's Unified Gauge-Based Analysis of Daily Precipitation over the Contiguous United States (CPC data). The CPC data has a continuous spatial coverage of 0.25 x 0.25 degrees at daily temporal resolution from 1948 to the present. The data are derived from over 13,000 daily rain gauges reporting across the contiguous U.S. Quality control procedures including ground-based radar precipitation and satellite corrections from the GOES 8 and GOES 10 IR data are used for data correction. Thirty-nine gridded stations fit within the Adirondack state park boundaries.

SST data were obtained from the Group for High Resolution Sea Surface Temperature (GHRSST) global Level 4 SST analysis. This product produces a daily 0.25 degree x 0.25 degree grid of SST at the NOAA National Climatic Data Center and uses optimum interpolation (OI) data from the 4-km AVHRR Pathfinder Version 5 time series. The data are bias adjusted using in-situ ship and buoy observations.

Both data sets cover the period from 1981–2010, and are converted into monthly anomaly time series. SST data were converted to a spatial resolution of 5 degrees. A total of 116 and 255 grids cover the northern Atlantic and Pacific basins, respectively. See Figure 4-1 for a map of the study area domains, and Table 4-1 for a summary list of the data sources used. The SST and precipitation data collected were used in the ANN forecasting model after the processing steps described in the following sections.

Table 4-1: Remote Sensing Data Products Summary.

| **Data Products** | **CPC Unified Gauge-Based Daily Precipitation over CONUS** | **GHRSST** |
|---|---|---|
| Data | Precipitation | SST |
| Primary Remote Sensing Instruments | Rain-gauge array | AVHRR |
| Agency | NOAA | NOAA |
| Temporal Resolution | Daily – converted to monthly anomaly | Daily – converted to monthly anomaly |
| Spatial Resolution | 0.25 x 0.25 degrees – converted to 5 x 5 degree | 0.25 x 0.25 degree |
| Ground-Truthing / Correction | Ground-based radar and satellite corrections from GOES 8 and GOES 10 IR data | Ship and buoy ground-truthing |
| Data Source | http://www.esrl.noaa.gov/psd/data/gridded/data.unified.html | http://podaac.jpl.nasa.gov/dataset/NCDC-L4LRblend-GLOB-AVHRR_OI |

### 4.1.2 Method

Long-term teleconnection analysis is hampered by using linear correlation to evaluate nonlinear systems (Franzke, 2009) and by an over-reliance on teleconnection indices that suffer from low signal to noise ratios. Additionally anthropogenic influences can affect local precipitation dynamics and mask weak signals of natural climate-based teleconnection (Mullon et al., 2013). This study aims to overcome these weaknesses by capturing the nonlinear and nonstationary teleconnection trends between Adirondack precipitation and SST in the Atlantic and Pacific basins. Teleconnection indices are generated to be site-specific, without relying on traditional published teleconnection regions in an attempt to capture the most significant oceanic regions to the study site, and are termed herein as non-leading teleconnections. Due to the strong evidence of leading teleconnection influence, it is expected that the non-leading teleconnection patterns developed herein will relate with the leading teleconnection oceanic regions regarding the ENSO, NAO and PNA described above. The non-leading teleconnections identified in this study will be derived independently and based exclusively on the data used herein.

The proposed analysis uses a stochastic technique for long-range forecasting of precipitation trends based on historic precipitation observation and teleconnection indices derived from SST observations. Local meteorological information associated with mesoscale and synoptic scale weather phenomena are not included in this approach, as they would be in a more deterministic weather prediction scheme for short-term meteorological analysis. Two time scales are used in this study, including a monthly and seasonal (i.e. 3 months) evaluation.

Monthly and seasonal time steps are used to balance and average the high spatiotemporal variability associated with weather events that cannot be explicitly detailed using this approach, as well as to compare the forecasting skill at various intervals. For the monthly analysis, data are calculated sequentially and continuously throughout all months. For the seasonal analysis, seasonal data are calculated for each season, where December-January-February (DJF) represents winter; March-April-May (MAM) represents spring; June-July-August (JJA) represents summer; and September-October-November (SON) represents fall.

An analytical framework for the study is provided in Figure 4-2. The framework includes the methodology used to identify the non-leading teleconnection patterns specific to the study area, as well as the ANN modeling used to forecast precipitation. As depicted in Figure 4-2, the first step is to obtain the precipitation and SST data, which are the hydrologic time series of interest. This step is followed by a wavelet transformation, where the original time series values are decomposed, using wavelets, into their fundamental signals and used to extract the scaled average wavelet power (SAWP) within defined frequency bands. Principal component analysis is then performed on the precipitation data array to create a single time series, representing the majority of the precipitation variability (WEOF-SAWP) within the Adirondacks. The new precipitation time series is then mapped, using pixel-wise correlation with the SST SAWP data to create the teleconnection map. Statistically significant teleconnection regions are then extracted from the teleconnection map and used as the exogenous data input into an ANN modeling framework. Historic precipitation values are also used as ANN input data. A fully

recurrent neural network model is then used to forecast precipitation one month or one season ahead.

As a comparison to the ANN modeling created as part of this research, a statistical downscaling model, using the SDSM software, is used to generate precipitation prediction across the same study site, using the same precipitation data observations. The ANN and SDSM models are compared using standard statistical metrics.

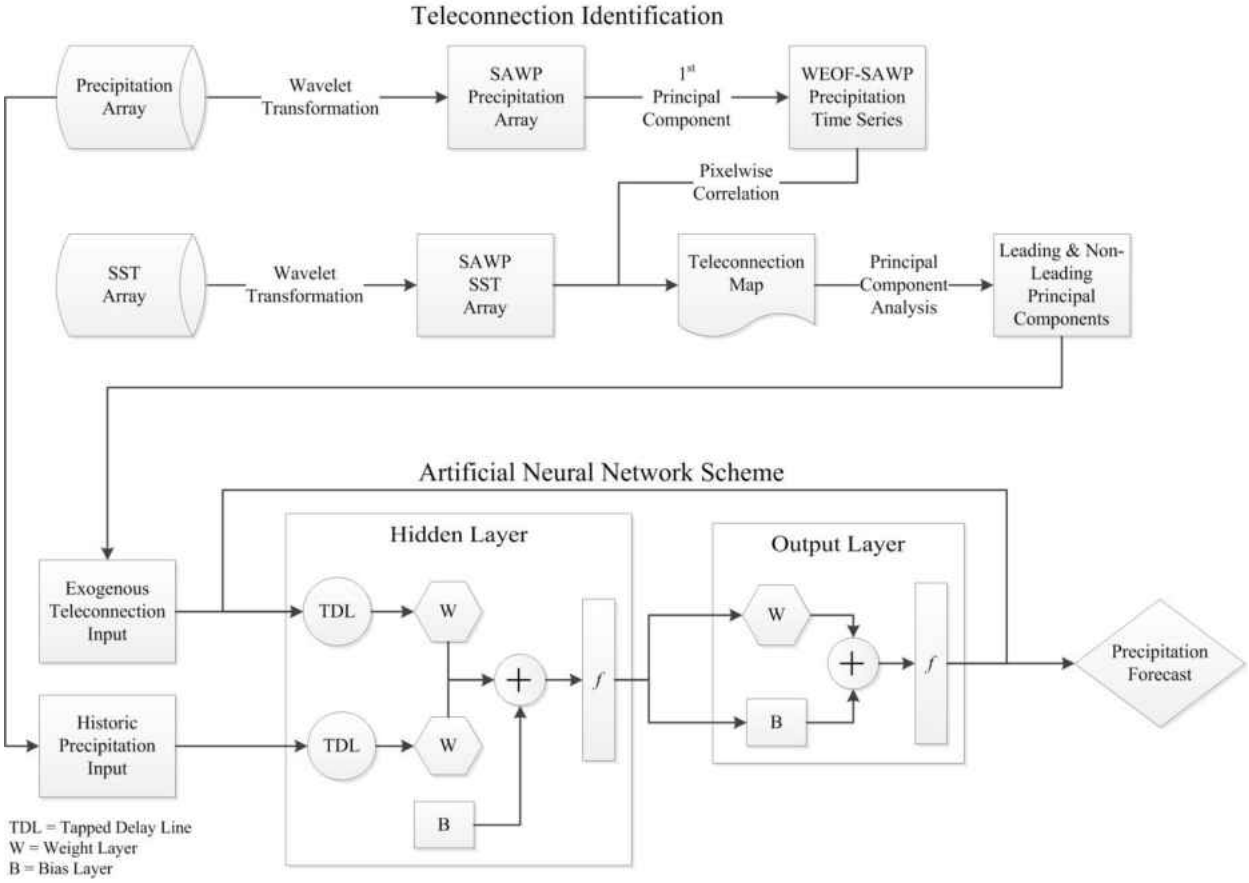


Figure 4-2: The analytical framework of the Long-Term study.

## 4.2 Wavelet Analysis Methodology

As discussed in the previous chapters, wavelet analysis has become a very useful tool for analyzing nonlinear and nonstationary behavior in environmental data. Traditional approaches, including Fourier analysis, were successfully used for stationary data analysis but were not well suited at describing nonstationarity. To solve this, Windowed Fourier Transform was developed for nonstationary signal processing, but is insufficient when analyzing environmental processes where a wide range of dominant frequencies are expected (Torrence and Compo, 1995). Wavelet analysis provides good time resolution for lower-period components as well as good period resolution for high-period components of the same signal (Anctil and Tape, 2004).

Wavelet transform is the process of decomposing time series data into an associated time-frequency domain. A type of spectral analysis, wavelet decomposition allows the determination of dominant localized variations of power (i.e. where the variance of the time series is largest for a given frequency) (Keener et al., 2010). By decomposing a time series into the time-frequency domain, one can determine both the dominant modes of variability and how those modes vary in time (Torrence and Compo, 1995). For this study, wavelet analysis was used to isolate dominant modes of North Atlantic and Pacific SST variability and to correlate these modes with like precipitation modes over the Adirondack state park to identify leading and non-leading teleconnection regions. Decomposed SST data are then used in an ANN precipitation forecasting model, herein denoted as WPC-ANN.

### 4.2.1 Scale-Averaged Wavelet Power (SAWP)

The continuous wavelet transform produces a wavelet power spectrum, a matrix representing the energy coefficients of the decomposed time series. The magnitude of the wavelet spectrum coefficients shows how well the wavelet matches the time series, with each scale depicting the amplitude of the time series. Individual scales are useful for demonstrating dominant modes of variability (Mwale et al., 2004). However, a range of dominant scales can be averaged into a SAWP and can be used to visualize the modulation of many frequencies within a single time series. The SAWP is the weighted sum of the wavelet power spectrum over scales $s_1$ and $s_2$:

$$\overline{W}_n^2 = \frac{\delta j \delta t}{C_\delta} \sum_{j=j_s}^{j_2} \frac{\left|W_n(s_j)\right|^2}{s_j} \quad \textbf{(8)}$$

Where $C_\delta$ is the reconstruction of a function from its wavelet. For this study the Morlet wavelet was used with a $C_\delta$ constant of 0.776.

### 4.2.2 Teleconnection Wavelet Processing

For the precipitation and SST datasets, two parallel procedures were carried out based on the time scales, monthly and seasonally. For each time scale, anomaly data were determined by calculating the monthly (or seasonal) averages of each time scale and subtracting that value from each gridded time series value, which represented the 30-year time

series for each gridded location. Each monthly (or seasonal) anomaly time series was then decomposed using the CWT process discussed in Chapter 3.

Figure 4-3 displays an example precipitation monthly time series result. The top image displays the inputted time series record (monthly anomaly). The central image shows the local wavelet power spectra where the dark red areas illustrate regions, in both time and frequency, where high wavelet power exists and blue areas illustrate regions of little power. The black bold lines are regions of statistical significance at the 95% confidence. The inverted arch shape represents the cone of influence in the image, which is determined by the length of the time series. In order for a wavelet to determine the power at each period, the non-dimensional wavelet expands to encompass larger periods and contracts to encompass smaller ones. Thus, larger periods suffer from "edge effects" where the larger wavelets have a much smaller window over which they can completely propagate through the time series. Therefore, zeroes are "padded" to either end for the time series to compensate. These regions of the graph outside of this line are not fully reliable because the wavelet is influenced by artificial zero values. Longer periods have a very narrow region of the graph where the data are considered valid, due to the edge effects. The graph on the right illustrates the global wavelet power. The bottom image represents the SAWP, which was calculated with the $s_1$ and $s_2$ scales set at 0.25 and 1 years, respectively.

After interpreting the wavelet power spectra for each SST and precipitation grids, a fairly consistent monthly trend emerged where dominant wavelet power was concentrated between the 0.25-1 year period bands. This is further evidenced by reviewing the global

wavelet power spectra on the right, which shows a dominant intrannual signal. This suggests that several dominant frequencies within this period band (i.e. the seasonal to annual period) exist that influences the SST and precipitation variability. Lower frequency wavelet power, illustrated by the dark red areas, are also present in the 1-8 year period as having strong wavelet power, but were generally shown to be less significant than the intrannual signal.

To efficiently capture the dominant frequencies, SAWP was calculated with the scales set at 0.25 and 1 year for all gridded monthly data. The gridded time series data are therefore converted from the original monthly anomaly time series into a new wavelet variability time series between the 0.25 and 1 year period band. This data are termed the precipitation and SST SAWP time series. The SAWP band window of 0.25–1 year was used for all gridded precipitation and SST datasets.

Figure 4-4 shows the same wavelet decomposition analysis, but using a seasonal SST grid (note that the center wavelet power spectra image uses contour lines for greater clarity for the seasonal data). Like the monthly power spectrum image in Figure 4-1, the 1-8 year period does display strong wavelet power, but in the seasonal analysis this period is statistically significant. The significance of the 1-8 year period in the seasonal data could suggest that the monthly time series data may contain noise that dampens the power spectrum at lower frequencies. For the seasonal analysis, the SAWP between the 1-8 year power spectrum is calculated. It is noted that the higher-frequency seasonal power spectrum is not able to be calculated as it is in the monthly analysis, because the data itself are seasonal.

The long-term seasonal wavelet findings are generally consistent with the findings in short-term investigation Chapter 3, where the interannual signal was shown to be dominant. The long-term monthly findings, showing dominant intrannual strength, differ somewhat from the short-term assessment.

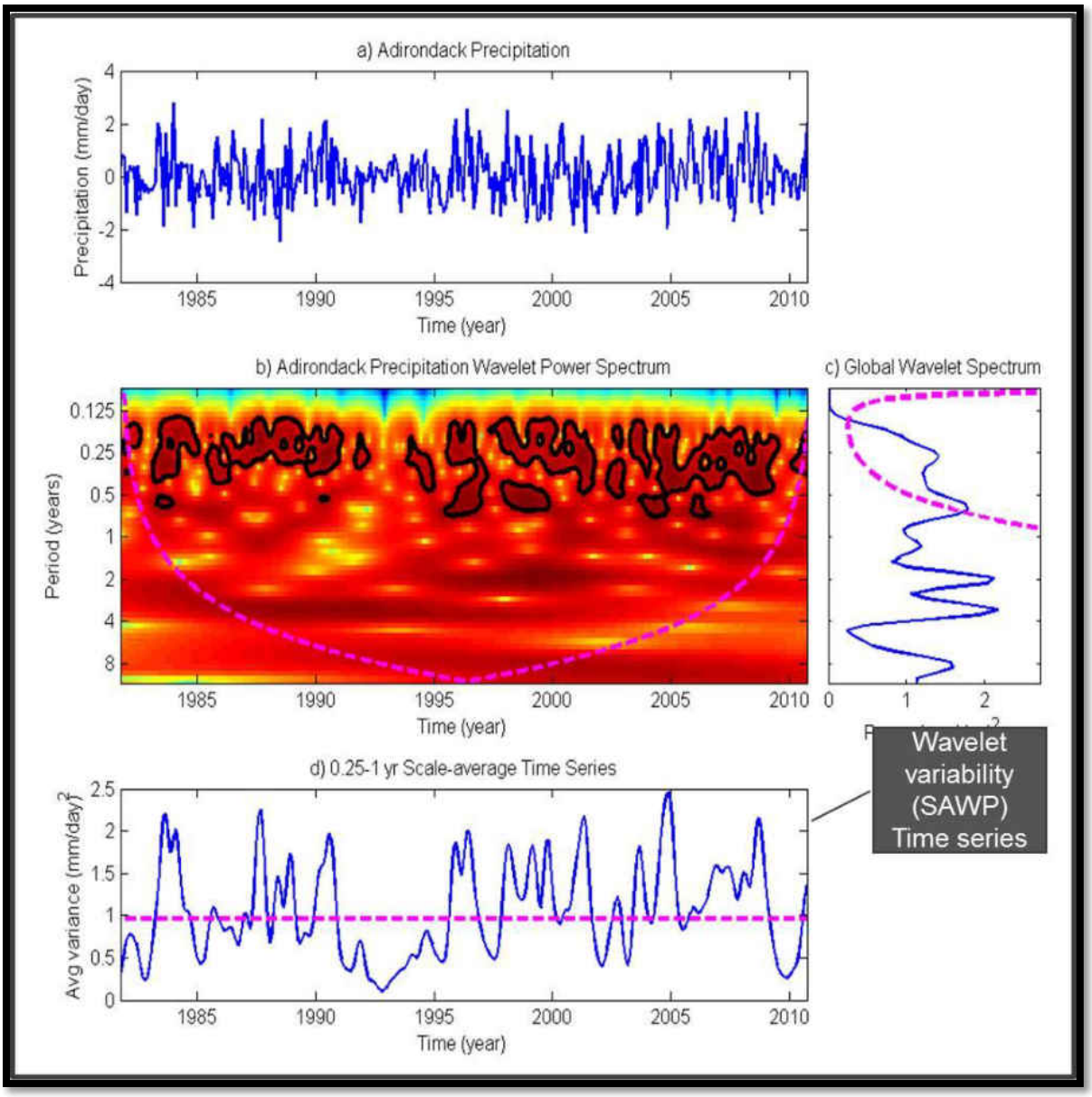


Figure 4-3: Wavelet decomposition example of a single monthly precipitation grid. The top image shows the monthly anomaly data with the X and Y axes representing time and SST in degrees, respectively. The middle image shows the expanded two-dimensional time frequency wavelet power spectra contour graph, with the X and Y axes representing time and period in years, respectively. The right graph shows the global wavelet power, with X and Y axes representing wavelet power in degrees C and period in years, respectively. The bottom image represents the SAWP time series based on the 1-8 year period band, with X and Y axes representing time in years and average variance, respectively.

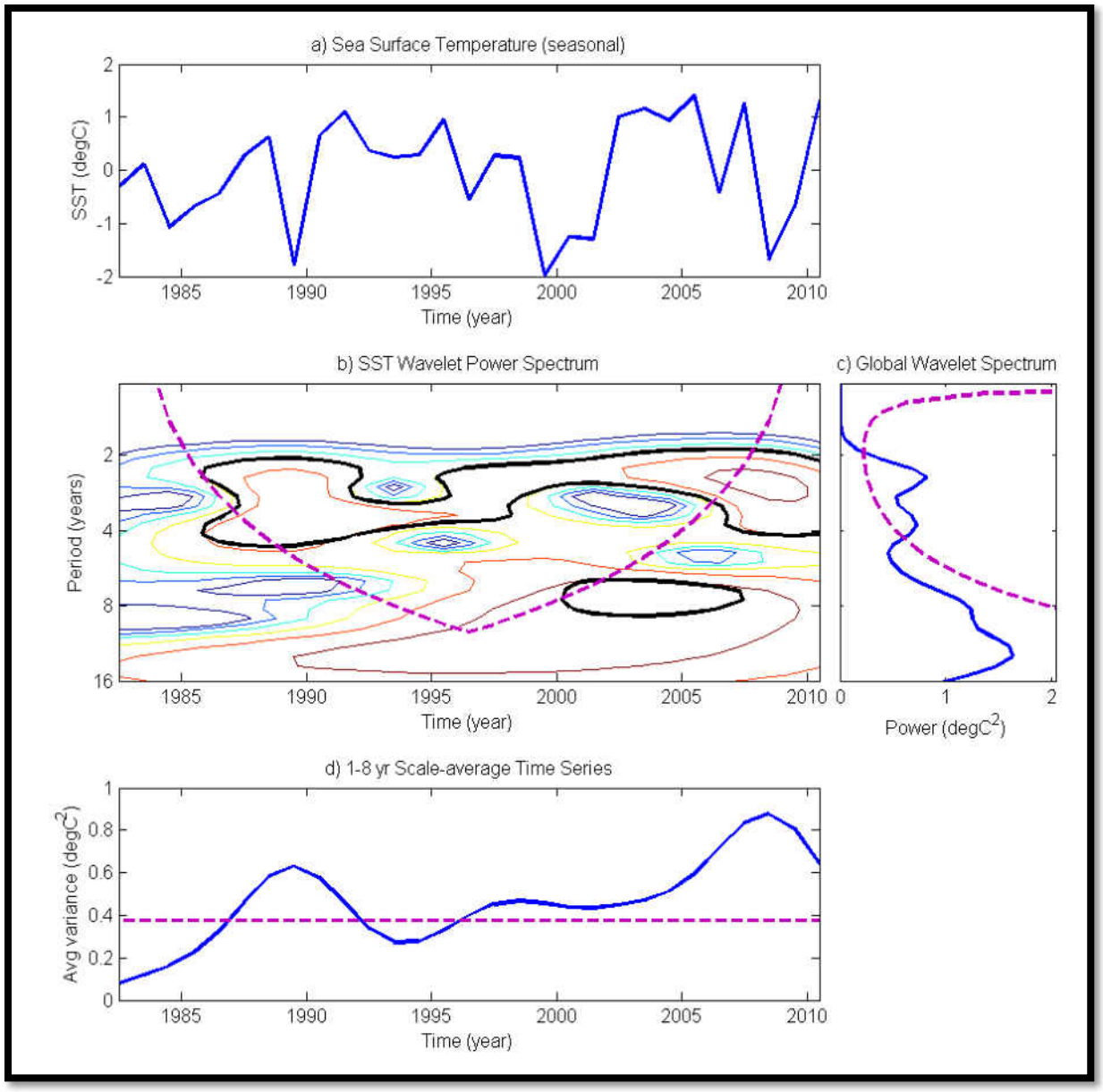


Figure 4-4: Wavelet decomposition example of a single seasonal SST grid.

### 4.2.3 Multivariate Principal Component Wavelet Processing

To establish specific oceanic teleconnection regions between the Adirondack precipitation and SST, pixel-wise correlation analysis was employed as a useful tool to identify regional teleconnection "hotspots." Due to the 39 precipitation grids within Adirondack, it is not practical to compare how each individual grid correlates with the SST within the oceanic boundaries. Principal component analysis was therefore used to consolidate the 39 gridded precipitation SAWP data into a single time series.

Principal component analysis uses empirical orthogonal functions (EOF) to transform multivariate data into spatially uncorrelated modes of variability (Venegas et al., 1996). The resulting PCs are time series data of the same length as the original and represent the variability within multivariate data. For large amounts of data with limited variability, a very few number of principal components (or just the first component) is needed to represent the majority of the variability within the original data.

The monthly SAWP precipitation time series data were consolidated into a single principal component (PC) time series. Calculating PCs on SAWP wavelet time series data has previously been denoted by Mwale et al. (2004) as Wavelet Principal Components (WPCs), and will be denoted as such herein. The resulting WPCs are considered frequency-compacted energy variability, and indicate the magnitude of events of time series averaged from the individual scales (Mwale et al, 2004).

WPCs for the monthly precipitation dataset were calculated. Because the precipitation within Adirondack region is relatively homogenous, the first WPC (i.e. WPC 1) was able to

describe approximately 90% of the overall variability, and is therefore a good representation of the entire dataset. The WPC precipitation dataset is plotted in Figure 4-5.

WPCs for the seasonal precipitation dataset were also calculated. WPCs for each season are shown in Figure 4-6. This figure shows that the strongest wavelet variability occurs in the last decade of the study.

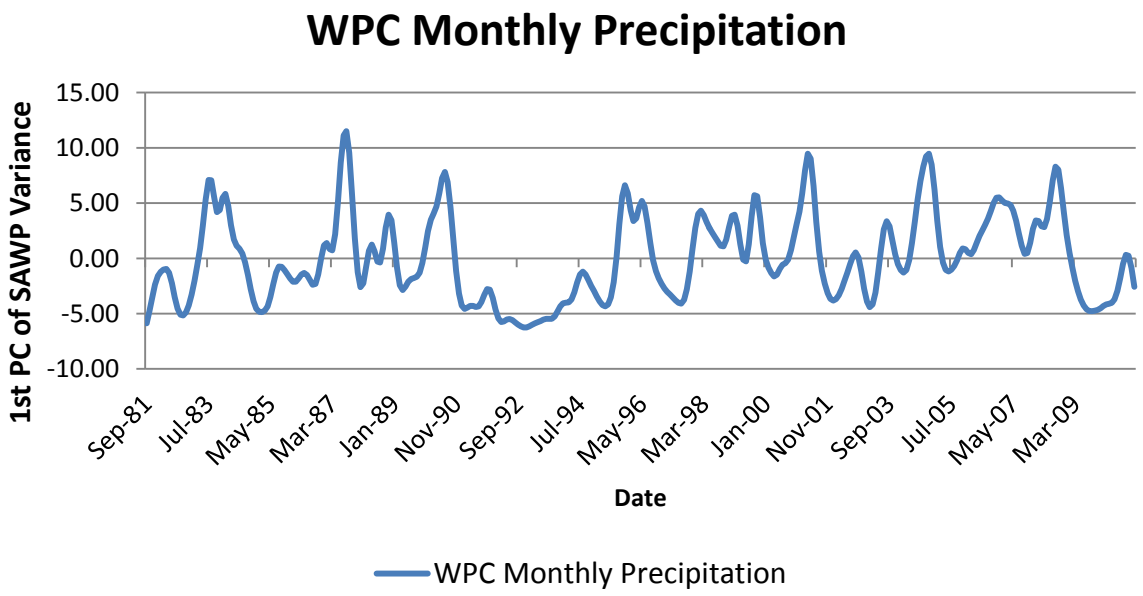


Figure 4-5: First Principal Component of the average precipitation SAWP time series (WPC).

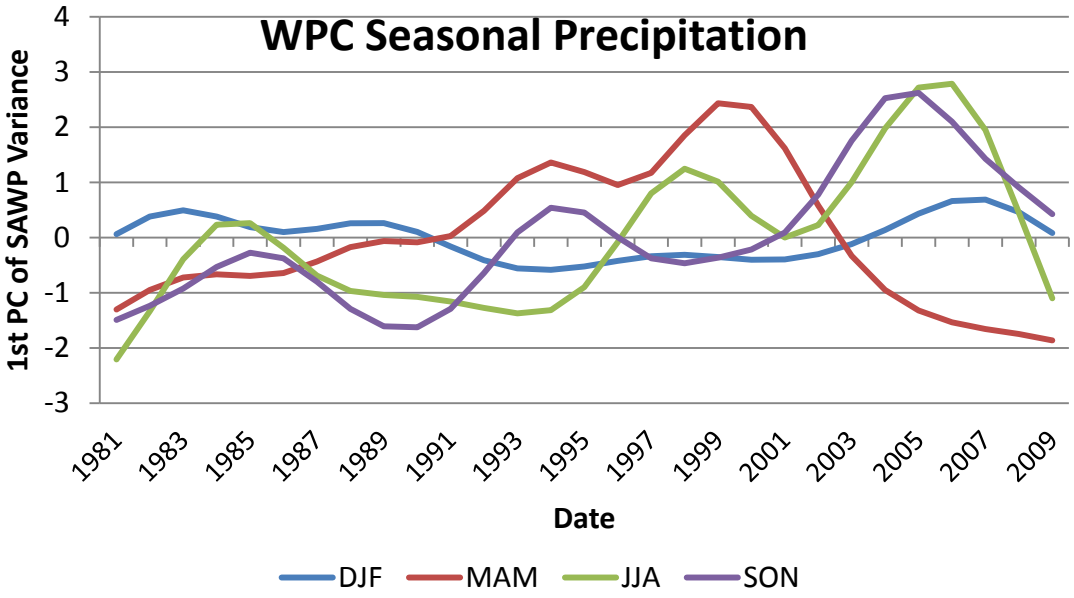


Figure 4-6: First Principal Component of the average precipitation SAWP seasonal time series (WPC).

### 4.2.4 Mapping Teleconnection Regions

#### 4.2.4.1 Monthly Teleconnection Mapping

For the monthly time scale, pixel-wise correlation maps were calculated between the WPC precipitation and the SAWP SST dataset at various lags from 1–12 months. This mapping is similar to the method used in Chapter 3, but uses the variability trends in lieu of simple anomalous data to capture the nonlinearity of potential teleconnection regions. Figure 4-7 shows the results of the teleconnection map based on the 12-month precipitation WPC lag, which was the strongest teleconnection map created from all lag periods. The colored contours represent correlation values with yellow and red indicating statistically significant positive

correlation at the 95% and 99% confidence interval, respectively, and light blue and dark blue indicating negative correlation at the 95% and 99% confidence intervals.

In the Atlantic basin, the strongest correlation zones are strongly oriented in the known NAO regions. The NAO is typically associated with high-pressure systems centered near the Azores and lower-pressure systems northward. This phenomenon is caused by the westerly atmospheric jet stream and the oceanic Gulf Stream. A signature of the NAO is an atmospheric pressure difference that causes a layered, meriodonal variability affect. The strong negative correlation from 45°-60°N and a positive correlation from 15°-45° are consistent with this NAO pattern. In the Pacific region, other known teleconnection patterns appear dominant. The strongest correlation occurs south of Alaska as well as along the eastern and western equatorial region, which are associated with the PNA and ENSO regions, respectively. Other leading teleconnection regions, including the Niño 1+2 and the Western Pacific Pattern (WPP) also appear. The strong adherence to traditional teleconnection regions between the Adirondacks and the northern Atlantic and Pacific basins reflects the historical evidence of the leading teleconnections with the northeastern U.S. discussed in Chapter 1, as well as the teleconnection maps prepared in Chapter 3.

In addition to the recreated leading teleconnection regions, high-correlation regions also exist that are not part of known teleconnection areas. A strong negative correlation off the western coast of Baja California (Index 6 in Chapter 3), as well as regions in the western North Atlantic, are present and are generally consistent with teleconnection patterns observed in Chapter 3, for northeastern teleconnections. These potential non-leading teleconnection

regions justify the need to examine teleconnection signals that relate specifically to a location, and not just rely on the traditional teleconnection indices.

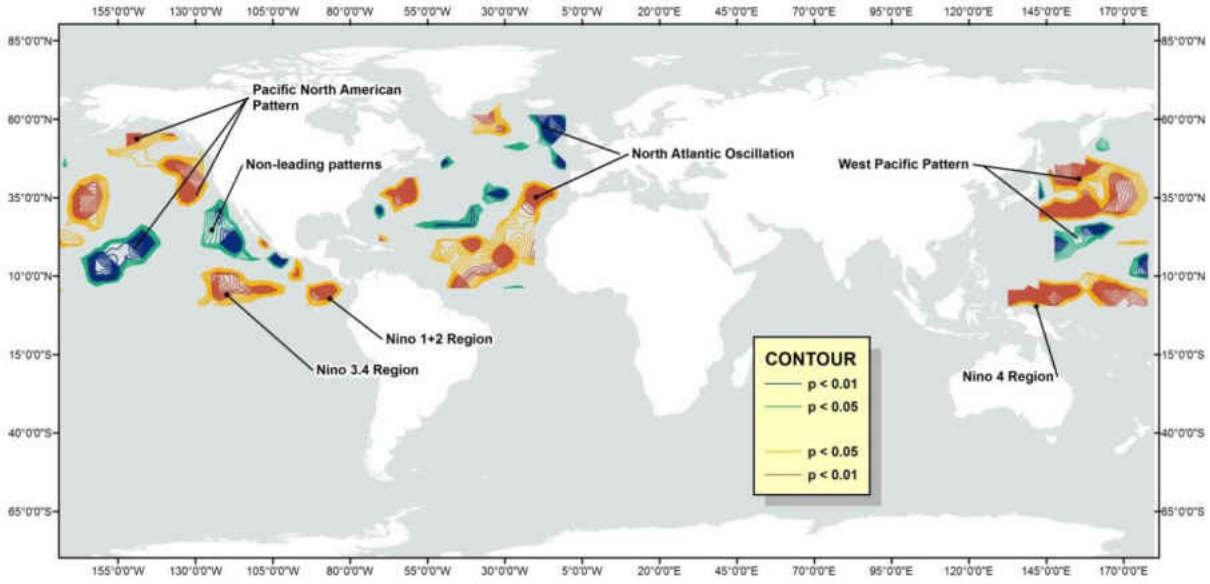


Figure 4-7: Results of the pixel-wise correlations between Adirondack precipitation WPC and SST SAWP time series. Yellow and dark red contours illustrate correlation at the 95 and 99% confidence intervals, respectively. Light and dark blue regions illustrate the negative correlation at the 95% and 99% confidence intervals, respectively.

From the 12-month lag map, the SAWP SST regions with greater than two statistically significant contiguous grids were extracted to create the teleconnection datasets, totaling 11 teleconnection regions. As with the precipitation data, principal component analysis was used to extract the underlying SAWP variability into single time series records for each extracted region. For each teleconnection region the first PC was calculated and extracted. The extracted teleconnection PCs are herein known as TPCs.

#### 4.2.4.2 Seasonal Teleconnection Mapping

For the seasonal teleconnection mapping, eight maps were produced, corresponding to the lagged seasonal relationships one and two seasons ahead between Adirondack precipitation WPC and the SST SAWP time series; see Figure 4-8 for an example. From the maps, all of the statistically significant SAWP grids were extracted to create individual teleconnection datasets. As with the monthly teleconnection mapping, the dominant statistical regions generally aligned with known teleconnections of the NAO, PNA and ENSO, however other potential non-leading site are also identified. The first 5 PCs from both the Atlantic and Pacific seasonal teleconnection datasets were then extracted. The extracted PCs are used as the exogenous input into the WPC-ANN model used for seasonal prediction.

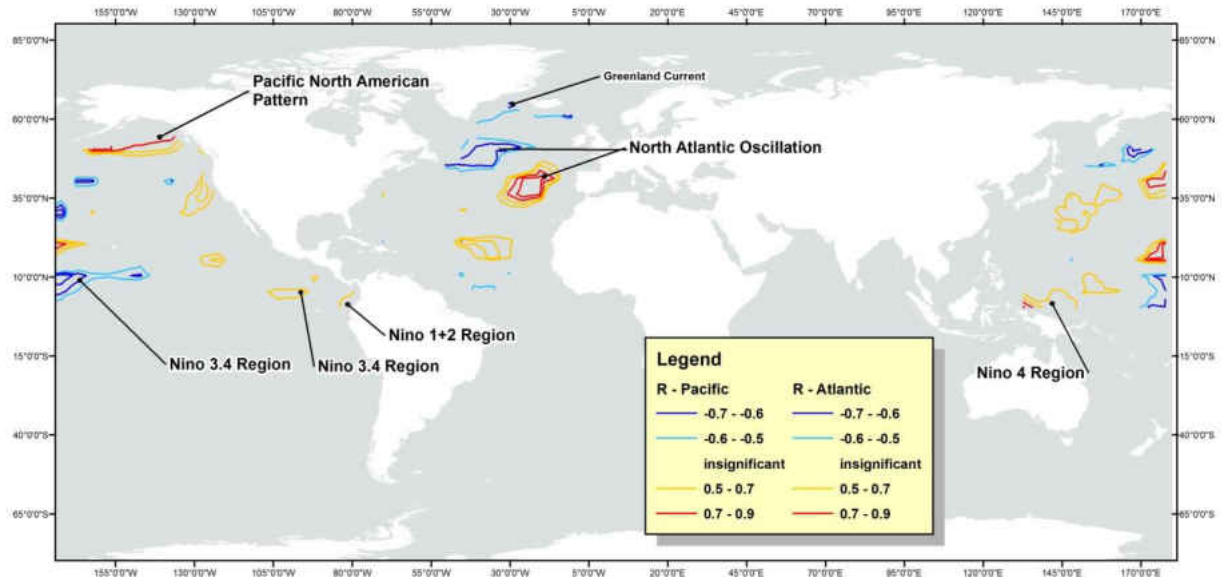


Figure 4-8: Results of the seasonal pixel-wise correlation map between the SON SST-SAWP dataset and the subsequent MAM precipitation WPC time series (i.e. two seasons ahead). Yellow and dark red contours illustrate correlation at the 95 and 99% confidence intervals, respectively. Light and dark blue regions illustrate the negative correlation at the 95% and 99% confidence intervals, respectively.

## 4.3 WPC-ANN Model

### 4.3.1 Monthly Precipitation Model

For this study a traditional 3-layer ANN model was used, consisting of an input, hidden and output layer. In these networks, the input layer receives the input variables used for estimating the output. This layer is the means of transferring data into the ANN model. The output layer includes the values predicted by the neural network, which represents the model output. The middle, hidden layer contains the number of nodes (neurons), which are predetermined by the modeler. Nodes within neighboring layers of the network are connected by links, with all nodes connected. Weights are assigned to each link to represent the connection strength of two nodes at both ends, which predicts the input-output relationship. Data enters the input layer, is processed through the hidden middle layer and is output into the output layer in the classic feed-forward architecture.

The basic feed-forward structure, however, is not sufficient for modeling the spatiotemporal variability of hydrologic processes, which requires that the neural network estimation procedure to be dynamic in order to more accurately interpret nonlinear feedback processes (American Society of Civil Engineers, 2000). To introduce memory into the static network architecture, a fully recurrent feedback process is also introduced that encloses the layers of the network, creating interconnection between the internal nodes. The defining equation for the model is as follows:

$$y(t) = f(y(t-1), y(t-2), ..., y(t-n_y), u(t-1), u(t-2), ...\mathrm{u(t-n}_u)) \quad (9)$$

Where the next value of the dependent output signal *y(t)* is regressed on the previous values of the output signal and previous values of an independent (exogenous) input signal for modeling nonlinear dynamic systems (Beale and Hudson, 2013).

For the monthly precipitation prediction model, the input variables are the raw, monthly, individual precipitation time series with the 11 TPCs serving as the independent input signal. For back-propagation, the Levenberg-Marquardt algorithm, a second-order nonlinear optimization technique, is used.

The goal of ANN modeling is to learn the underlying patterns and relationship from a subset of input data, from which the model is trained. To test the capabilities of the network, an unknown input (input not used in the training step) is used for prediction purposes. This process is known as generalization. Standard generalization approaches split the input data into three components, a training set used to train the neural net, the validation set to test the performance of the neural net that are not used in the trained set, and finally a test set for checking the overall performance. Training should be stopped when the minimum validation error occurs. When training is not stopped, overtraining can occur and the performance of the test set decreases, despite improvements on the training data. This approach can have drawbacks when dealing with limited time series data, as three division sets require less data to be used in the original training step. In nonlinear systems, it is important to train neural networks on a representative set of data. In limited nonlinear time series such as precipitation, low-frequency recurrent signals might not be “seen” in the training set if the time series is not

long enough. It is therefore important to have the longest training set possible without compromising the test set.

As a check against potential overfitting within the proposed WPC-ANN model, the ANN capacity test developed by Wang et al. (2014) was used. In this test, the training samples are measured against a theoretical lower and upper bound of the capacity of an ANN, which is based on the number of inputs, the number of outputs and the number of neurons used in the model. The algorithm is presented below:

$$L \le P \le L\left(1 + {}^{N}\!/_{M}\right) \tag{10}$$

Where *P* is the capacity of the ANN (number of samples), *L* is the number of neurons, *N* is the number of inputs and *M* is the number of outputs in the ANN model. For the proposed WPC-ANN model, the number of samples represents the number of forecasted months (30), the number of neurons used is 10, the number of inputs is 12, and the number of outputs is 1. Thus, for the proposed model, $10 \le 30 \le 130$ meets the capacity test.

Because the precipitation dataset encompasses a spatial coverage area of 39 gridded time series with relative homogeneity, it was not necessary to create 39 different precipitation forecast models. Instead, the average raw precipitation value was calculated across the entire precipitation dataset. This average precipitation time series was used as the historical precipitation input into the WPC-ANN model. The 11 TPCs extracted from the teleconnection mapping were used as the exogenous input into the WPC-ANN model.

To evaluate the performance of the WPC-ANN model, Pearson's correlation coefficient and root mean square error (RMSE) values were calculated. After the optimum model was developed using the average precipitation data, the individual gridded precipitation data were used to forecast the precipitation at each grid one month ahead. Maps were produced between the observed precipitation events and the predicted values for each grid. The prediction statistics are calculated as follows:

Pearson's Correlation:

$$r = \left( \sum_{i}^{n} \frac{\left[ \left( x_i - \mu(x) \right) \times \left( y_i - \mu(y) \right) \right]}{\sqrt{\sum_{i}^{n} \left( x_i - \mu(x) \right)} \times \sqrt{\sum_{i}^{n} \left( y_i - \mu(y) \right)^2} } \right) \quad \textbf{(11)}$$

Where *r* is the correlation coefficient, *n* is the number of observations in a time series, *x* is a measurement of precipitation dataset time series, *y* is the predicted measurement and $\mu$ is the mean.

Root Mean Square Error:

$$RMSE = \left[ \frac{1}{n} \sum_{k=1}^{n} \left( x_k - y_k \right)^2 \right]^{\frac{1}{2}} \quad \textbf{(12)}$$

For this study, training was conducted over the period from 1981 – 2007 and testing was carried out from 2008–2010, a period of 30 months, or approximately 8.3% of the time series record.

### 4.3.2 Seasonal Prediction Model

The seasonal precipitation model is similar to the monthly precipitation model, except the input variables are the raw, seasonal, individual precipitation time series with the 10 seasonal TPCs serving as the independent input signal. For back-propagation, the Levenberg-Marquardt algorithm, a second-order nonlinear optimization technique, is used.

To evaluate the performance of the WPC-ANN model, Pearson's correlation coefficient and root mean square error (RMSE) values were calculated for each gridded time series. Maps were produced for each seasonal result between the observed precipitation events and the predicted values for each grid. For the seasonal model, training was conducted over the period from 1981–2003 and testing was carried out from 2004–2010, a period of 7 years, or approximately 25% of the time series record. Because the seasonal prediction model is based on such a limited data record (only 7 forecast seasons), Bayesian regularization is used for the neural network generalization, which is a generalization scheme optimized for small datasets. Bayesian regularization allows the data set to be split simply into a training and validation set, and does not require the capacity test to be performed.

## 4.4 Statistical Downscaling Comparison

As a comparison to the proposed WPC-ANN model, a statistical downscaling approach for forecasting Adirondack precipitation using the Statistical Downscaling Model (SDSM) software (version 4.2) was used. Statistical downscaling is a technique that uses large-scale weather and climate output data as predictor information and calibrates the information with local observed precipitation to downscale the data resolution for use at regional and local

scales. The large-scale weather and climate data are obtained from observed data, or from the NCEP-NCAR re-analysis information. Long-term forecasts are made possible by substituting general circulation model (GCM) predictor data output over the period of interest. The SDSM model was only used as a comparison against the monthly WPC-ANN model, due to data limitations of the seasonal time series record.

Statistical downscaling is similar to model output statistics and perfect prog analyses used for hydrologic time series prediction, and is a popular climate forecasting technique because of its low computational cost and fast assessments of local climate phenomena, primarily temperature and precipitation. The SDSM software is described as a hybrid of the stochastic weather generator and transfer function downscaling techniques, as large-scale circulation patterns and atmospheric moisture variables are used to condition local-scale weather generator parameters, and stochastic techniques are used to downscale GCM time series to correlate with local observations (Wilby and Dawson, 2007).

SDSM has been used to forecast long-term precipitation trends globally, including areas in the northeastern United States and southeastern Canada, with similar precipitation and climatic patterns. Tryhorn and DeGaetano (2011) used SDSM to calibrate rain gage data over much of the northeastern U.S., including the Adirondacks, using NCEP-NCAR re-analysis data and the HADCM3 GCM for long-term climate forecasts of extreme rainfall events. In southeastern Canada, north of the Adirondacks region, SDSM-based precipitation forecasting was found to be more capable than an ANN model in reproducing various statistical characteristics of observed data in its downscaled results (Khan et al., 2006). This study,

however, relied on NCEP-NCAR re-analysis data to calibrate and forecast precipitation from 1961–2000, which is the available NCEP-NCAR data record for the SDSM software, and did not use GCM data for precipitation forecasting.

### 4.4.1 SDSM Methodology

SDSM requires an input of the predicted (observed precipitation), predictors (observed or NCEP-NCAR re-analysis climate variables, such as temperature, surface pressure and wind direction) for conditioning, and GCM output of the same predictors for long-term climate forecasting. For this study, the NCEP-NCAR reanalysis climate data from 1981–2001 was used as well as the HADCM3 GCM output for climate forecasting from 2002–2010, the list of climate variables available is provided in Table 4-2 below. A screening procedure is available for the modeler to analyze which predictor variables to use in the downscaling scheme. Available statistics include correlation matrices, partial correlation analysis and scatterplots.

Table 4-2: List of predictor variables available for the SDSM downscaling using NCEP reanalysis data and HADCM3 GCM output data. Bold variables were used for downscaling and forecasting precipitation at the Adirondacks.

| Variable | NCEP | HADCM3 |
|---|---|---|
| Mean Sea Level Pressure | X | X |
| Surface Airflow Strength | X | X |
| **Surface Zonal Velocity** | **X** | **X** |
| **Surface Meriodonal Velocity** | **X** | **X** |
| **Surface Vorticity** | **X** | **X** |
| **Surface Wind direction** | **X** | **X** |
| Surface Divergence | X | X |
| 500 hPa Airflow Strength | X | X |
| 500 hPa Zonal Velocity | X | X |

| Variable | NCEP | HADCM3 |
|---|---|---|
| 500 hPa Meridional Velocity | X | X |
| 500 hPa Vorticity | X | X |
| **500 hPa geopotential Height** | **X** | **X** |
| 500 hPa Wind Direction | X | X |
| 500 hPa Divergence | X | X |
| 850 hPa Airflow Strength | X | X |
| 850 hPa Zonal Velocity | X | X |
| 850 hPa Meridional Velocity | X | X |
| 850 hPa Vorticity | X | X |
| 850 Geopotential Height | X | X |
| 850 hPa Wind Direction | X | X |
| 850 hPa Divergence | X | X |
| Specific Humidity at 500 hPa | X | X |
| Specific Humidity at 850 hPa | X | X |
| **Surface Specific Humidity** | **X** | **X** |
| **Mean Temperature at 2m** | **X** | **X** |

For this study, surface zonal velocity, surface meriodonal velocity, surface vorticity, surface wind direction, 500 hPa geopotential height, surface specific humidity and mean temperature were chosen as predictor variables. A correlation matrix of all predictand and predictor variables was performed to measure the various explanatory power of the chosen variables. The results of the correlation matrix are provided in Table 4-3. From this table, surface vorticity has the strongest correlation with observed precipitation. Surface specific humidity and mean temperature at 2m are also relatively strong as compared with the remaining variables, which demonstrate weak explanatory power.

Table 4-3: Correlation matrix for all variables, including predictand (precipitation) and predictors for the period of 1980-2001.

| | **Observed Precip** | **Surface Zonal Velocity** | **Surface Merid. Velocity** | **Surface Vorticity** | **Surface Wind Direction** | **500 hPa Geopotential Height** | **Surface Specific Humidity** | **Mean Temp at 2m** |
|---|---|---|---|---|---|---|---|---|
| Observed Precipitation | 1 | 0.017 | 0.019 | 0.487 | 0.011 | 0.004 | 0.197 | 0.132 |
| Surface Zonal Velocity | 0.017 | 1 | 0.12 | 0.071 | -0.668 | 0.004 | 0.143 | 0.141 |
| Surface Meriodonal Velocity | 0.019 | 0.12 | 1 | 0.074 | -0.211 | 0.406 | 0.288 | 0.302 |
| Surface Vorticity | 0.487 | 0.071 | 0.074 | 1 | -0.08 | 0.13 | 0.37 | 0.369 |
| Surface Wind Direction | 0.011 | -0.668 | -0.211 | -0.08 | 1 | -0.122 | -0.21 | -0.196 |
| 500 hPa Geopotential Height | 0.004 | 0.004 | 0.406 | 0.13 | -0.122 | 1 | 0.808 | 0.868 |
| Surface Specific Humidity | 0.197 | 0.143 | 0.288 | 0.37 | -0.21 | 0.808 | 1 | 0.932 |
| Mean Temp at 2m | 0.132 | 0.141 | 0.302 | 0.369 | -0.196 | 0.868 | 0.932 | 1 |

Calibration of the predictors and predictand is then carried out by computing the parameters of multiple regression equations via optimization algorithms (dual simplex or ordinary least squares). A Weather Generator step then creates ensembles of synthetic weather series given the predictor atmospheric variables (NCEP-NCAR). This step enables the verification of calibrated models and the creation of artificial time series for present climate conditions. In the final step, the Scenario Generator creates ensembles of synthetic weather time series given the predictor variables supplied by the GCM (HADCM3) for past, present or

future climate conditions, instead of the observed (NCEP) predictors. For this study, the HADCM3 data were downscaled for the period of 2001–2010 and used as a comparison forecast tool to the WPC-ANN forecasting model. Bias correction for the proposed precipitation downscaling was modeled with a value of 1.0, indicating no bias correction, as is the recommended default value in the SDSM user's manual.

Statistical downscaling has a natural advantage over the WPC-ANN forecasting model, as the former can account for the weather variables presented in Table 4-2. These variables are based on observed and modeled climate patterns, which offer a more comprehensive view of the regional climate processes than the WPC-ANN model. Additionally, the SDSM tool enables ensembles of climate scenarios to be used, which permits the development of a risk and uncertainty analysis.

Weaknesses also exist using statistical downscaling that the WPC-ANN model avoids. In statistical downscaling, predictor data are dependent on the underlying GCM, meaning that GCM regions that may not explain complex climate processes, such as those due to poor boundary forcings will not produce suitable downscaling results. In addition, the regression techniques used during the model calibration process can have poor performance when the predictor-predictand relationships are nonlinear and nonstationary. Low-frequency climate variability, such as periodic climate teleconnection patterns, is also known to be problematic with respect to statistical downscaling, as the process does not account for the memory effect of the ocean-atmospheric system.

## 4.5 Results and Discussion

### 4.5.1 Monthly WPC-ANN & SDSM Comparison

The results from the monthly WPC-ANN and SDSM precipitation forecast models are provided in Figure 4-9. The top graph represents the average precipitation across the 39 precipitation grids within the Adirondack park boundary. The middle graph depicts the SDSM statistical downscaling precipitation forecasting model overlain in dark red with the average precipitation, beginning in 2001. The bottom graph depicts the WPC-ANN precipitation forecasting model, with the red graph representing the training and validation through region through 2008 and the green representing the forecasting, through 2010. Statistics of the results are provided in Table 4-4.

The SDSM forecasting period is longer than the WPC-ANN forecast period because the available NCEP reanalysis data used to condition the SDSM model ended in 2001. SDSM was able to reasonably capture the seasonality of the precipitation, but was not able to match the peaks and lows of the precipitation trends. This could be a result of poor correlation statistics between the available climate variables and the precipitation observed within the Adirondack State Park during the weather conditioning process. Because the model depends on the relationships between the predictor variables and the predictand, it is unlikely that highly accurate predictions can be made. Another potential source of error comes from the nonstationary and nonlinear behavior of the precipitation data, as well as a reliance on the HADCM3 GCM output for this region. Statistical downscaling is known to perform poorly when using highly nonstationary and nonlinear time series or when the underlying GCM large-scale

weather variables are not properly calibrated with the trained data. It is expected that other climate predictands, such as temperature, would perform better than precipitation using similar downscaling approaches.

The WPC-ANN model was developed using two primary inputs, the historic precipitation as well as the 11 TPCs, identified using wavelet analysis, during the teleconnection mapping. Both inputs were fed into the WPC-ANN to produce precipitation forecasts one month ahead. As seen in Figure 4-6, the WPC-ANN model produced a stronger precipitation forecast than SDSM, with correlation values approximately double that of the SDSM forecasting model. Additionally, the WPC-ANN model more accurately predicted the summer and autumn highs and winter and spring low precipitation amounts that occur over the Adirondack region. Multiple WPC-ANN models were developed with different forecasting periods (not shown). As the forecasting duration increased above 30 time steps, the predictive capabilities of the WPC-ANN model decreased rapidly. This is expected in ANN models, as the model is highly sensitive to previous precipitation conditions and generally performs better with longer time series records. Thus, longer precipitation time series record may produce stronger forecasting capabilities.

The WPC-ANN model was developed using the average historic precipitation across all 39 grids. To determine the spatiotemporal variability within the Adirondack region, the same WPC-ANN model was applied to each of the individual precipitation gridded time series, thus forecasting 39 separate precipitation values. Figure 4-10 shows the WPC-ANN model results distributed across the Adirondack park boundary, with the left image showing the correlation

contours between the observed and forecasted precipitation, and the right image showing the average precipitation error contours. Interestingly, stronger correlation trends are observed throughout most of the Adirondack region as compared to the average precipitation graph shown in Figure 4-9, with correlation values of 0.8 achieved in the northeastern quadrant. The forecast error image on the right however shows that some regions of high correlation areas in the northeast are significantly under predicting the precipitation, suggesting that some high correlation regions are somewhat misleading, and not producing as strong of a forecast as suggested. Generally, the central portions of the Adirondacks have a small average forecast error and a relatively strong correlation with the observed prediction trends. The northeast and southeast quadrant with higher average precipitation errors could be due to topographical conditions, such as the mountainous northeast, which may produce anomalous precipitation trends difficult to forecast.

Table 4-4: Statistics for the SDSM and WPC-ANN precipitation forecasting models.

| Statistic | SDSM | WPC-ANN |
|---|---|---|
| Pearson's r | 0.3 | 0.6 |
| RMSE | 1.3 | 0.7 |

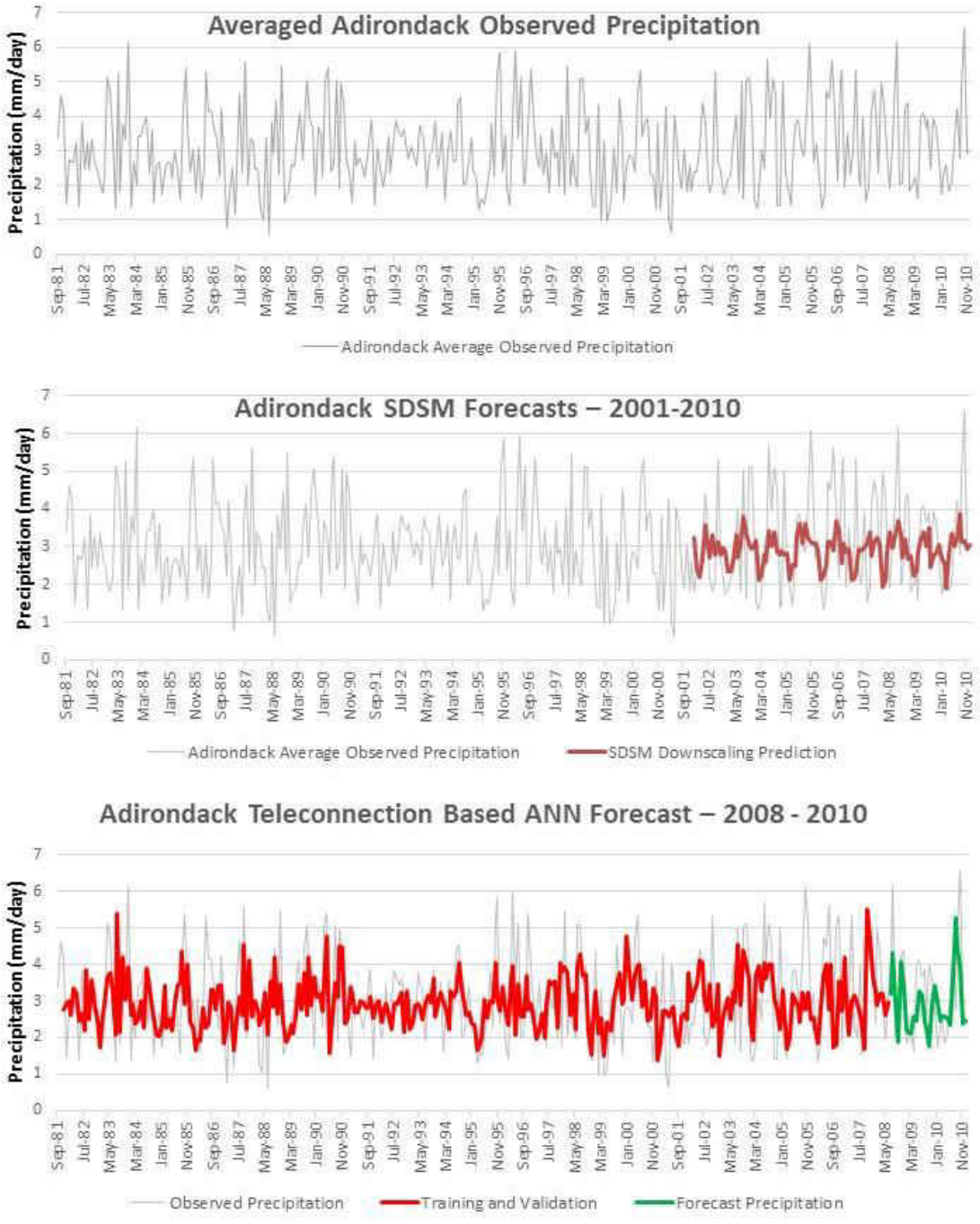


Figure 4-9: Forecasting results for the SDSM and WPC-ANN models. The top graph shows the original averaged precipitation time series for the Adirondack State park in gray. The middle graph shows the SDSM model precipitation forecasting results from 2001 – 2010 overlain in dark red. The bottom image graph shows the WPC-ANN model overlain with the red indicating the training and validation and green showing the precipitation forecasting

Table 4-5: WPC-ANN Forecasting Tabular Results.

| Date | Average Observed Precipitation (mm/day) | WPC-ANN Forecasting Result (mm/day) | Error (mm/day) |
|---|---|---|---|
| Jul-08 | 4.22 | 3.18 | 1.04 |
| Aug-08 | 6.14 | 4.30 | 1.85 |
| Sep-08 | 2.02 | 3.00 | -0.98 |
| Oct-08 | 2.08 | 1.88 | 0.20 |
| Nov-08 | 4.22 | 4.07 | 0.15 |
| Dec-08 | 4.40 | 3.40 | 1.00 |
| Jan-09 | 1.83 | 2.20 | -0.37 |
| Feb-09 | 2.00 | 2.12 | -0.12 |
| Mar-09 | 2.25 | 2.12 | 0.14 |
| Apr-09 | 1.61 | 2.56 | -0.95 |
| May-09 | 3.95 | 2.45 | 1.51 |
| Jun-09 | 4.10 | 3.29 | 0.82 |
| Jul-09 | 3.64 | 3.16 | 0.49 |
| Aug-09 | 3.90 | 2.43 | 1.46 |
| Sep-09 | 2.48 | 2.25 | 0.23 |
| Oct-09 | 3.94 | 1.77 | 2.17 |
| Nov-09 | 3.67 | 2.89 | 0.77 |
| Dec-09 | 2.76 | 3.43 | -0.68 |
| Jan-10 | 1.75 | 2.97 | -1.22 |
| Feb-10 | 2.27 | 2.48 | -0.20 |
| Mar-10 | 2.61 | 2.58 | 0.03 |
| Apr-10 | 1.84 | 2.58 | -0.74 |
| May-10 | 2.06 | 2.53 | -0.46 |
| Jun-10 | 3.30 | 2.33 | 0.97 |
| Jul-10 | 4.24 | 3.66 | 0.58 |
| Aug-10 | 2.78 | 5.28 | -2.50 |
| Sep-10 | 4.83 | 4.51 | 0.33 |
| Oct-10 | 6.58 | 3.93 | 2.65 |
| Nov-10 | 2.94 | 2.38 | 0.56 |
| Dec-10 | 3.02 | 2.44 | 0.57 |

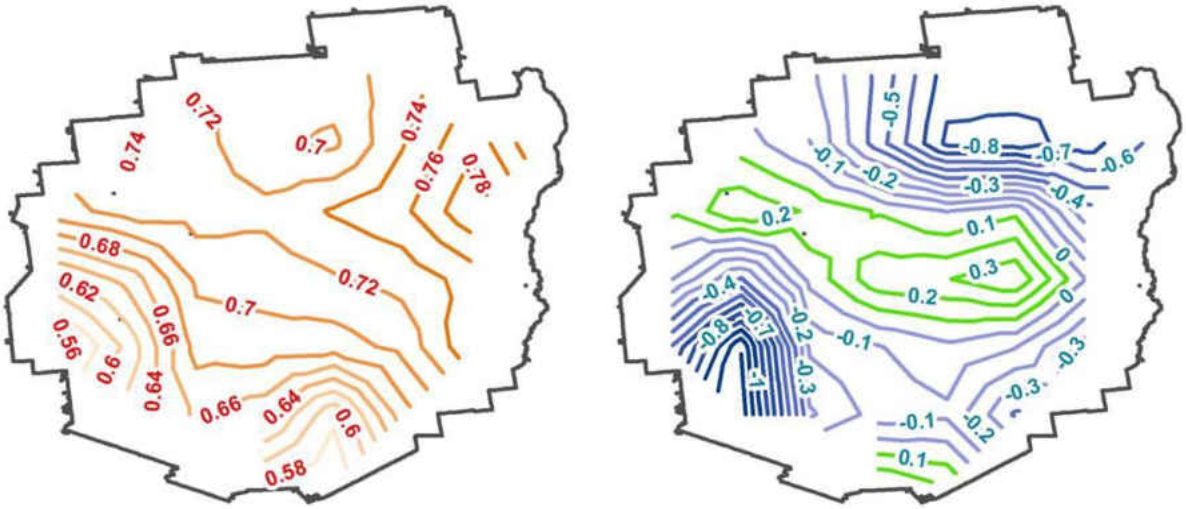


Figure 4-10: WPC-ANN modeled precipitation forecasting applied to all 39 Adirondack precipitation grids. The image on the left shows the correlation contour. The image on the right shows the precipitation error with blue indicating under prediction and green indicating over prediction.

### 4.5.2 Seasonal WPC-ANN Results

Pearson's correlation and RMSE maps for the DJF predicted Adirondack precipitation are shown on Figure 4-11. Maps 4-11A and 4-11B show the WNN results where all 10 TPCs for both ocean regions were included, whereas maps 4-11C and 4-11D only include one ocean region's TPC input from the Pacific and Atlantic basins, respectively. From map 4-11A, correlation values exceeding 0.8 are realized in the southern two thirds of the Adirondack state park with lower predictive values in the northern reaches, suggesting high predictive ability for winter rainfall patterns in most of the study area. RMSE values (map 4-11B) show a similar pattern with lower values of 0.4 occurring in the south, suggesting greater predictive skill than the north. Map 4-

11C shows a much lower predictive skill than Map 4-11D, suggesting that the Atlantic basin is more important to winter Adirondack precipitation than the Pacific.

Regions of low correlation in Figure 4-11A do not necessarily represent specific areas within Adirondack that are not well predicted. As illustrated in Figure 4-12, the 2006 winter season was very accurately predicted, with an $R^2$ value 0.83 across all 30 gridded regions. Similar predictive skill was found for the 2004 winter period, whereas 2010 produced a prediction correlation value of only 0.43. These findings suggest that strong winter predictability exists for Adirondack region, which could be reasonably relied upon for watershed-scale precipitation forecasts.

The predictive ability for the seasons MAM, JJA and SON did not produce meaningful predictive skill. A longer time series record is expected to improve the WNN model forecasts for the remaining seasons. Additional measures, including higher temporal resolution, could also benefit the WNN model, as the higher frequency variability would be captured in the wavelet decomposition process and ultimately used in the WNN model.

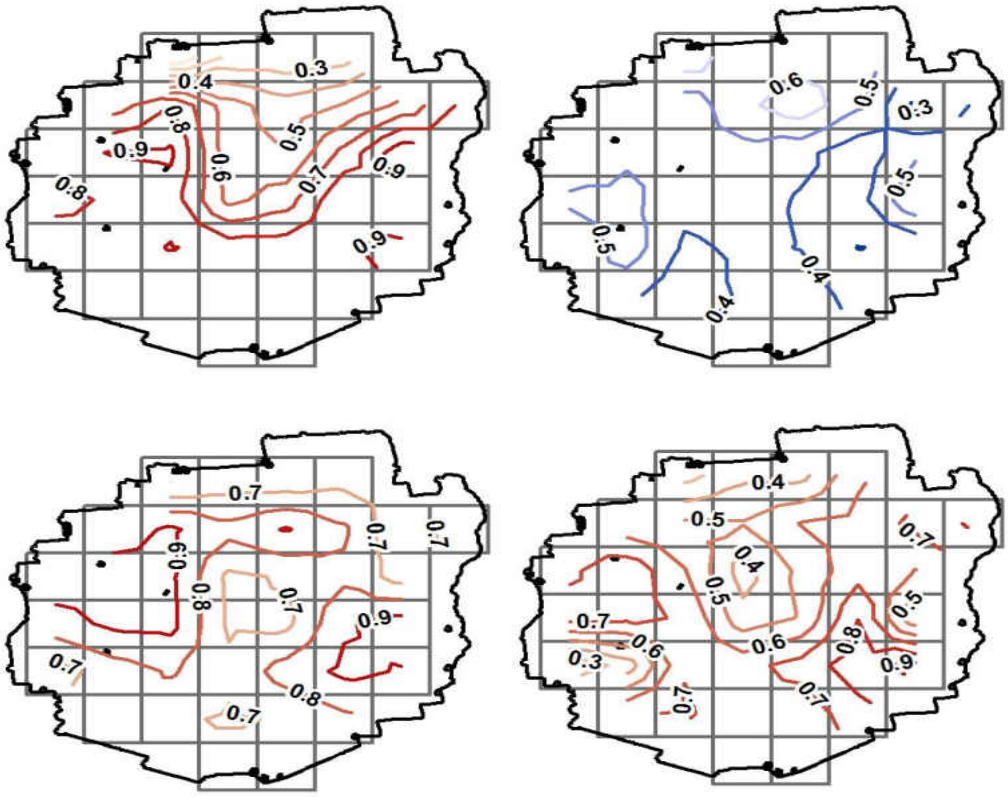


Figure 4-11: Maps A-D in clockwise order from the upper left. For all maps, the bold black outline represents the Adirondack state park boundaries, with the gray squares representing the 39 precipitation data grids. Map A is the correlation (red contours) between observed and predicted precipitation from 2004-2010 for each grid using all 10 TPCs from the Atlantic and Pacific oceans. Map B is the same as A except uses the RMSE (blue contours) statistic. Maps C and D are the same as map A except only the TPCs from the Pacific (Map C) or Atlantic (Map D) are included in the WNN model, respectively.

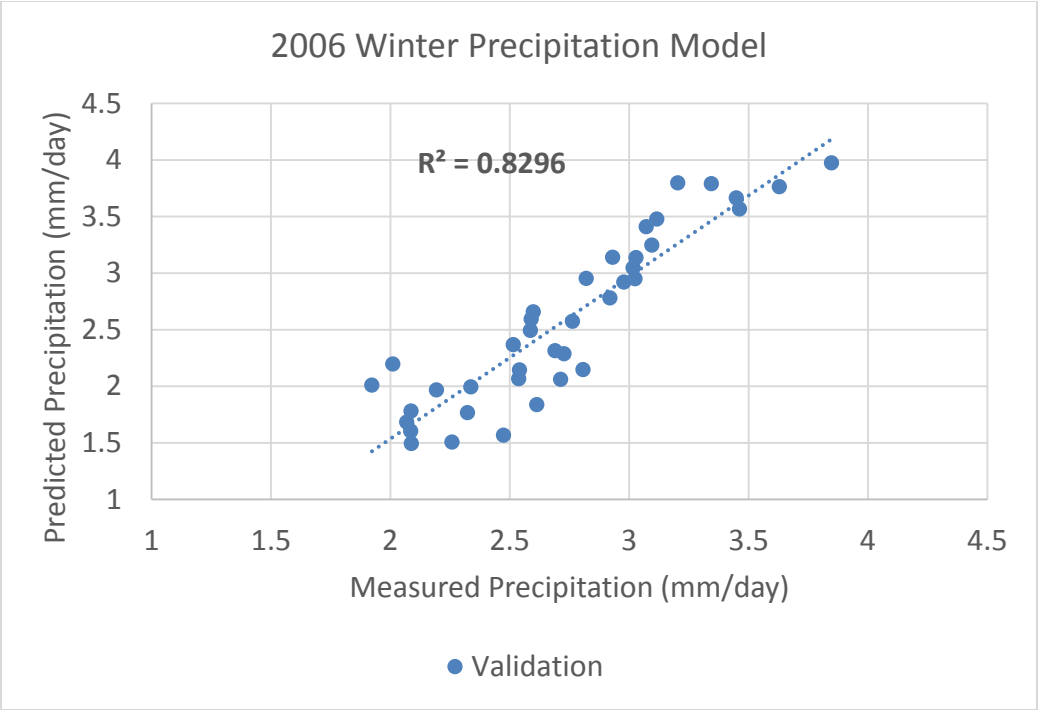


Figure 4-12: Seasonal precipitation prediction skill for winter 2006 year for all 39 precipitation grids.

From the three science questions posed, it has been demonstrated that wavelet-based teleconnection mapping may be a suitable process for deriving the leading teleconnection regions associated with the Adirondacks that have been established in the literature, including the ENSO, NAO and PNA oceanic regions, using a 12-month lag. Non-leading teleconnections areas, such off the west coast of Baja California, as well as the other leading patterns unknown to be associated with northeastern U.S. climate, like the Western Pacific Pattern, are similarly identified in this process. These potential non-leading patterns warrant detailed analysis, as they may play an important role in the underlying processes that drive regional climate patterns in the northeastern U.S. Using only historic precipitation, as well as the TPCs identified

in the teleconnection mapping, a WPC-ANN precipitation forecasting model has been developed to predict precipitation amounts one month and one season ahead. When considering seasonal forecasts, the winter period exhibits the strongest predictive skill. The monthly WPC-ANN prediction model was also compared against a statistical downscaling approach using the SDSM software. The WPC-ANN precipitation model was able better predict precipitation trends than the SDSM model.

# CHAPTER 5 : CONCLUSION

## 5.1 Short-Term Investigation Conclusion

High-resolution remote-sensing imagery provides us with the global ability to identify climate teleconnection signals in terms of greenness and precipitation variability in local areas of the northeast U.S. Using nonlinear spectral wavelet analysis, a consistent biennial to triennial low-frequency signal can be collectively identified among greenness, precipitation and SST anomalies in the Northern Atlantic and Pacific oceans. Linear correlation "hot spots," specifically within the 0–3 month time lag for the Atlantic and eastern equatorial Pacific basin, and 8-9 months (or longer) for the central and western equatorial Pacific basin proved useful in locating specific ocean regions where teleconnection signals may exist. Extracted indices reproduce those areas of the oceans where traditional teleconnections exist, including the NAO and ENSO. This method of linear correlation is effective in screening and identifying these links over a vast area.

Such an integrated remote sensing and wavelet analysis establishes a strong correlation link between each terrestrial site and surrounding oceans, with $R^2$ values up to 0.50 for greenness and 0.42 for precipitation when considering multi-ocean influences using lagged stepwise regression analysis. This analytical process suggests that linear models using simple SST anomalies could be used for prediction estimates. Because SST indices are extracted from optimal locations, often within known teleconnection regions, this method may prove more

successful for localized studies than those that consider published teleconnection data suitable for broader areas.

Wavelet analysis reveals signals of climate teleconnection between 16 and 30 months across all datasets, illustrating a clear biennial to triennial variability between SST and terrestrial response. A 3–9 month lag between equatorial Pacific SST and precipitation variability suggests that Pacific-induced Atlantic SST anomalies were present, which truly represent a teleconnection influence through tropical Pacific–Atlantic processes. Negative greenness patterns following La Niña winters also suggest that ENSO was influencing mid-latitude jet stream patterns. For precipitation patterns, NAO indices explain the most variability, with the Index 2 region showing a consistent 2-month lead time over each study site. This finding is consistent with earlier studies, suggesting that NAO-induced precipitation patterns were associated with nor'easter storm tracking following jet stream fluctuations.

Precipitation was not a strong source of greenness variability when high periods of precipitation were associated with low periods of greenness and vice versa. Temperature variations might have stronger association with greenness variability. This allows ENSO teleconnections to have a significant effect on northeastern greenness by modulating the jet stream patterns, which can generate wildly different temperature anomalies as polar weather is forced into and out of the region.

Finally, this analysis considered a 10-year short-term time series sufficiently long to establish small interannual teleconnection signals; however, the datasets are not long enough to investigate longer-term interannual or decadal periods. Caution must be used when drawing

conclusions from short-climate records, which may lead to the misinterpretation of statistical noise from deterministic causes. A longer time series, additional site locations, and climate variables are needed to further explore the teleconnection processes and the ecosystem response.

## 5.2 Long-Term Teleconnection Summary

Using wavelet theory, objectively identified teleconnection regions associated with the Adirondack state park in the northeastern U.S. have been used to forecast monthly and seasonal precipitation trends using a new artificial neural network scheme. Monthly wavelet decomposition reveals that the dominant wavelet power period for both precipitation and SST domains is the seasonal to annual frequency band. Lower interannual periods from 1-8 years exhibit high wavelet energy but are statistically insignificant. Seasonal wavelet decomposition reveals similar lower interannual periods of high wavelet power from 1-8 years, which are statistically significant. The lack of statistical significance for the interannual period in the monthly data differs from the results in the short-term study for the SPW site. Considering that the Adirondacks cover a much larger area than just the SPW, and that the study time period is 30 year, and not 10, these statistical differences can largely be explained.

Using the dominant wavelet energy band of 0.25–1 year for monthly time scale, and 1-8 years for the seasonal time scale, the SAWP was calculated to extract the wavelet variability within the frequency bands. Using principal component analysis, the precipitation SAWP dataset was consolidated into its first principal component (WPC of precipitation).

Teleconnection mapping was performed using lagged pixel-wise correlation between the WPC precipitation and the SAWP SST domain.

The identified monthly teleconnection patterns in the 12-month lag as well as the 6-month lag for seasonal teleconnection maps are consistent with the leading teleconnections defined in the literature as having influence in the northeastern U.S. Potential non-leading teleconnections associated with the northern Atlantic and Pacific oceans are also visible. Principal component analysis is again used to extract the underlying teleconnection variability across the SST domain, known as the TPCs.

Eleven TPCs for the monthly analysis and 10 TPCs for the seasonal analysis were used as inputs into a neural network model (WPC-ANN). Analysis reveals that the monthly WPC-ANN model performed better than a statistical downscaling model using the SDSM software, with the WPC-ANN model able to better model observed precipitation from summer and autumn highs and winter and spring lows. Average precipitation errors are concentrated in the northeastern and southeastern quadrants of the Adirondacks, suggesting that local topography may play a role in forecasting quality. It is noted that previous studies demonstrated stronger SDSM forecasting skill in the northeast U.S. and southeastern Canada, but were evaluated at different time domains, longer (40 year) data period, and did not include GCM output (Khan et al., 2006; Tryhorn and DeGaetano, 2011).

For the seasonal WPC-ANN model, only the winter prediction model produced meaningful results. For the year 2006, the precipitation model accurately forecasted precipitation across the entire 39 gridded Adirondacks region, with $R^2$ values of approximately

0.83. While the monthly forecast model reproduced overall more consistent precipitation estimates, the seasonal winter forecast model is shown to most accurately reproduce precipitation totals.

The methodology proposed herein uses a stochastic technique for long-range forecasting of precipitation trends based on historic precipitation observation and teleconnection indices derived from SST observations. Monthly and seasonal time steps are used for forecasting, consistent with similar teleconnection-based forecasting efforts. Mesoscale and synoptic scale weather patterns and climatic data are not explicitly considered, which would be needed for higher temporal resolution and deterministic-based forecasting. It is anticipated that a longer time series record as well as the incorporation of other predictor variables, such as temperature, SLP and wind direction will improve the long-range forecast skill, which is recommended for future research.

From the science questions posed, the leading and potential non-leading teleconnection patterns have been identified using wavelet-based teleconnection mapping. The unified leading and non-leading areas were successfully used in a WPC-ANN modeling scheme and compared against a statistical downscaling model. The WPC-ANN was able to better predict precipitation than the downscaling approach within the study region and time domain, consistent with the author's hypothesis. This study suggests that non-leading teleconnection regions may play an important role in various regional and local climatic processes. The ability to objectively determined teleconnection regions associated with a particular terrestrial area may also prove

an important means of developing predictive models, or to augment existing models for regional and local forecasting purposes.

## 5.3 Concluding Remarks

This research has demonstrated the effectiveness of using nonlinear spectral analysis for the identification of statistical teleconnection regions to the northeastern U.S. and the northern Atlantic and Pacific oceans. For this study, wavelet analysis was used as the principal tool for the identification and exploration of teleconnection patterns. Wavelet analysis techniques have become a popular method for analyzing environmental data, and for establishing causal relationships between multiple time series. It is noted, however, that wavelet analysis is not the only, nor necessarily the best, nonlinear statistical technique for teleconnection identification. A recent observation of wavelet-based research is the existence of wavelet "energy leakage" (Peng et al., 2009), which is a function of the wavelet band overlap degree and the time resolution for discrete wavelet transforms. Wavelet energy leakage could lead to weaker statistical relationships that may otherwise be present and dominant. Other methods, including Hilbert Huang Transform and Singular Value Decomposition are available to analysts for forecasting hydrologic time series. These methods have their own strengths and weaknesses, and may be more or less successful than wavelet analysis in revealing underlying statistical relationships, though such an analysis is beyond the scope of this research. Regardless of the method used, nonlinear spectral analysis with nonlinear ANN modeling offers a relatively inexpensive means for developing moderate to long-term climate forecasts. Such models are expected to become highly valuable tools for water managers, as well as for

industries that rely on the accurate forecast of precipitation patterns, such as hydropower producers, water supply utilities, agricultural management facilities, flood management authorities, urban planning (related to climate change forecasts), and water quality regulation.

## LIST OF REFERENCES


Abbot, J., Marohasy, J., 2012. Application of artificial neural networks to rainfall forecasting in Queensland, Australia. Adv. Atmos. Sci. 29, 717–730. doi:10.1007/s00376-012-1259-9

AIRMAP, 2004. El Nino, the North Atlantic Oscillatio nand New England Climate. Winter Teleconnections and Climate Prediction [WWW Document]. URL http://airmap.unh.edu/background/nao.html (accessed 5.14.12).

American Society of Civil Engineers, 2000. Artificial Neural Networks in Hydrology. I: Preliminary Concepts. J. Hydrol. 5, 115–123.

Anctil, F., Tape, D.G., 2004. An exploration of artificial neural network rainfall-runoff forecasting combined with wavelet decomposition. J. Environ. Eng. Sci. 3, S121–S128. doi:10.1139/s03-071

Athanasiadis, P.J., Ambaum, M.H.P., 2009. Linear Contributions of Different Time Scales to Teleconnectivity. J. Clim. 22, 3720–3728. doi:10.1175/2009JCLI2707.1

Barnston, A.G., Livezey, R.E., 1987. Classification, seasonality and persistency of low frequency atmosphereic cirulcation patterns. Mon. Weather Rev. 115, 1083–1126.

Beale, M., Hudson, M.T., 2013. Neral Network Toolbox Users Guide. Natick, Massachusetts.

Beckage, B., Osborne, B., Gavin, D.G., Pucko, C., Siccama, T., Perkins, T., 2008. A rapid upward shift of a forest ecotone during 40 years of warming in the Green Mountains of Vermont. Proc. Natl. Acad. Sci. U. S. A. 105, 4197–4202. doi:10.1073/pnas.0708921105

Buermann, W., 2003. Interannual covariability in Northern Hemisphere air temperatures and greenness associated with El Niño-Southern Oscillation and the Arctic Oscillation. J. Geophys. Res. 108, 4396. doi:10.1029/2002JD002630

Cai, W., Sullivan, A., Cowan, T., 2009. Rainfall teleconnections with indo-pacific variability in the WCRP CMIP3 models. J. Clim. 22, 5046–5071. doi:10.1175/2009JCLI2694.1

Cassou, C., Terray, L., Hurrell, J.W., Deser, C., 2004. North Atlantic winter climate regimes: Spatial asymmetry, stationarity with time, and oceanic forcing. J. Clim. 17, 1055–1068. doi:10.1175/1520-0442(2004)017<1055:NAWCRS>2.0.CO;2

Center, N.C.D., 2012. Southern Oscillation Index (SOI) [WWW Document]. URL http://www.ncdc.noaa.gov/teleconnections/enso/indicators/soi.php (accessed 5.14.12).

Cho, J., Yeh, P.J.-F., Lee, Y.-W., Kim, H., Oki, T., Kanae, S., Kim, W., Otsuki, K., 2010. A study on the relationship between Atlantic sea surface temperature and Amazonian greenness. Ecol. Inform. 5, 367–378. doi:10.1016/j.ecoinf.2010.05.005

Chuvieco, E., Huete, A., 2010. Fundamental of Satellite Remote Sensing. CEC Press, Boca Raton.

Coulibaly, P., Baldwin, C.K., 2005. Nonstationary hydrological time series forecasting using nonlinear dynamic methods. J. Hydrol. 307, 164–174. doi:10.1016/j.jhydrol.2004.10.008

Czaja, A., Frankignoul, C., 1999. Influence of the North Atlantic SST on the atmospheric circulation. Geophys. Res. Lett. 26, 2969–2972. doi:10.1029/1999GL900613

Dixon, P.G., Goodrich, G.B., Cooke, W.H., 2008. Using Teleconnections to Predict Wildfires in Mississippi. Mon. Weather Rev. 136, 2804–2811. doi:10.1175/2007MWR2297.1

Enfield, D.B., Mayer, D.A., 1997. Tropical Atlantic sea surface temperature variability and its relation to El Nino Southern Oscillation. J. Geophys. Res. 102, 929–945.

Enfield, D.B., Mestas-nu, A.M., Trimble, P.J., 2001. The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U . S . 28, 2077–2080.

Enfield, D.B., Mestas-Nuñez, A.M., Trimble, P.J., 2001. The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. Geophys. Res. Lett. 28, 2077–2080. doi:10.1029/2000GL012745

Farge, M., 1992. Wavelet Transforms And Their Applications To Turbulence. Annu. Rev. Fluid Mech. doi:10.1146/annurev.fluid.24.1.395

Foster, D.R., Motzkin, G., Slater, B., 1998. Land-Use History as Long-Term Broad-Scale Disturbance: Regional Forest Dynamics in Central New England. Ecosystems. doi:10.1007/s100219900008

Franzke, C., 2009. Nonlinear Processes in Geophysics Multi-scale analysis of teleconnection indices : climate noise and nonlinear trend analysis 65–76.

Franzke, C., Woollings, T., 2011. On the Persistence and Predictability Properties of North Atlantic Climate Variability. J. Clim. 24, 466–472. doi:10.1175/2010JCLI3739.1

Greene, C.H., 2012. The winter of our discontent. Sci. Am. 307, 50–55.

Grinsted, A., Moore, J.C., Jevrejeva, S., 2004. Nonlinear Processes in Geophysics Application of the cross wavelet transform and wavelet coherence to geophysical time series 561–566.

Hodson, D.L.R., Sutton, R.T., Cassou, C., Keenlyside, N., Okumura, Y., Zhou, T., 2009. Climate impacts of recent multidecadal changes in Atlantic Ocean Sea Surface Temperature: a multimodel comparison. Clim. Dyn. 34, 1041–1058. doi:10.1007/s00382-009-0571-2

Holman, I.P., Rivas-Casado, M., Bloomfield, J.P., Gurdak, J.J., 2011. Identifying non-stationary groundwater level response to North Atlantic ocean-atmosphere teleconnection patterns using wavelet coherence. Hydrogeol. J. 19, 1269–1278. doi:10.1007/s10040-011-0755-9

Hubeny, J.B., King, J.W., Reddin, M., 2011. Northeast US precipitation variability and North American climate teleconnections interpreted from late Holocene varved sediments. Proc. Natl. Acad. Sci. U. S. A. 108, 17895–900. doi:10.1073/pnas.1019301108

Huber, S., Fensholt, R., 2011. Analysis of teleconnections between AVHRR-based sea surface temperature and vegetation productivity in the semi-arid Sahel. Remote Sens. Environ. 115, 3276–3285. doi:10.1016/j.rse.2011.07.011

IPCC, I.P.O.C.C., 2007. Climate Change 2007 - The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC. Science (80-. ). 1009. doi:volume

Johnstone, J. a., 2010. A quasi-biennial signal in western US hydroclimate and its global teleconnections. Clim. Dyn. 36, 663–680. doi:10.1007/s00382-010-0755-9

Joseph, R., Nigam, S., 2006. ENSO evolution and teleconnections in IPCC's twentieth-century climate simulations: Realistic representation? J. Clim. 19, 4360–4377. doi:10.1175/JCLI3846.1

Keener, V.W., Feyereisen, G.W., Lall, U., Jones, J.W., Bosch, D.D., Lowrance, R., 2010. El-Niño/Southern Oscillation (ENSO) influences on monthly NO3 load and concentration, stream flow and precipitation in the Little River Watershed, Tifton, Georgia (GA). J. Hydrol. 381, 352–363. doi:10.1016/j.jhydrol.2009.12.008

Khan, M.S., Coulibaly, P., Dibike, Y., 2006. Uncertainty analysis of statistical downscaling methods. J. Hydrol. 319, 357–382. doi:10.1016/j.jhydrol.2005.06.035

Leathers, D., Yarnal, B., Palecki, M., 1991. The Pacific/North American teleconnection pattern and United States climate. Part I: Regional temperature and precipitation associations. J. Clim. 4, 517–528.

Los, S.O., Collatz, G.J., Bounoua, L., Sellers, P.J., Tucker, C.J., 2001. Global interannual variations in sea surface temperature and land surface vegetation, air temperature, and precipitation. J. Clim. 14, 1535–1549. doi:10.1175/1520-0442(2001)014<1535:GIVISS>2.0.CO;2

Maier, H.R., 1997. Determining Inputs for Neural Network Models of Multivariate Time Series 12, 353–368.

Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables : a review of modelling issues and applications 15, 101–124.

Marshall, J., Kushnir, Y., Battisti, D., Chang, P., Czaja, A., Dickson, R., Hurrell, J., McCartney, M., Saravanan, R., Visbeck, M., 2001. North Atlantic climate variability: Phenomena, impacts and mechanisms. Int. J. Climatol. 21, 1863–1898. doi:10.1002/joc.693

McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. 5, 115–133. doi:10.1007/BF02478259

Meehl, G.A., T.F. Stocker, W.D. Collins, P. Friedlingstein, A.T. Gaye, J.M. Gregory, A. Kitoh, R. Knutti, J.M. Murphy, A. Noda, S.C.B. Raper, I.G. Watterson, A.J.W. and Z.-C.Z., 2007. Global Climate Projections. In: Climate Change 2007:The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York, NY, USA.

Minns, a. W., Hall, M.J., 1996. Artificial neural networks as rainfall-runoff models. Hydrol. Sci. J. 41, 399–417. doi:10.1080/02626669609491511

Montroy, D., 1997. Linear Relation of Central and Eastern North American Precipitation to Tropical Pacific Sea Surface Temperature Anomalies *. J. Clim. 10, 541–558.

Mullon, L., Chang, N.-B., Jeffrey Yang, Y., Weiss, J., 2013. Integrated remote sensing and wavelet analyses for screening short-term teleconnection patterns in northeast America. J. Hydrol. 499, 247–264. doi:10.1016/j.jhydrol.2013.06.046

Mwale, D., Yew Gan, T., Shen, S.S.P., 2004. A new analysis of variability and predictability of seasonal rainfall of central southern Africa for 1950-94. Int. J. Climatol. 24, 1509–1530. doi:10.1002/joc.1062

National Weather service Climate Prediction Center, 2012. Pacific/North American (PNA). [WWW Document], n.d. URL http://www.cpc.ncep.noaa.gov/data/teledoc/pna.shtml (accessed 5.14.12).

New York State Department of Environmental Conservation, 2005. Siamese Ponds Wilderness and Dug Mountain, Forks Mountain and Chatiemac Primitive Areas. Albany.

Notaro, M., Wang, W.-C., Gong, W., 2006. Model and Observational Analysis of the Northeast U . S . Regional Climate and Its Relationship to the PNA and NAO Patterns during Early Winter 3479–3506.

Peng, Z.K., Jackson, M.R., Rongong, J. a., Chu, F.L., Parkin, R.M., 2009. On the energy leakage of discrete wavelet transform. Mech. Syst. Signal Process. 23, 330–343. doi:10.1016/j.ymssp.2008.05.014

Raible, C., 2003. Precipitation and Northern Hemisphere regimes. Atmos. Sci. Lett. 5, 43–55. doi:10.1016/j.atmoscilet.2003.12.001

Rao, Y.R.S., Krishna, B., 2009. Modelling hydrological time series data using wavelet neural network analysis.

Rial, J.A., Sr, R.A.P., Beniston, M., Noblet-ducoudré, N.D.E., Prinn, R., 2004. NONLINEARITIES , FEEDBACKS AND CRITICAL THRESHOLDS WITHIN THE EARTH ' S CLIMATE SYSTEM 11–38.

Sutton, R.T., Hodson, D.L.R., 2005. Atlantic Ocean forcing of North American and European summer climate. Science 309, 115–8. doi:10.1126/science.1109496

Sutton, R.T., Hodson, D.L.R., 2007. Climate Response to Basin-Scale Warming and Cooling of the North Atlantic Ocean. J. Clim. 20, 891–907. doi:10.1175/JCLI4038.1

Tang, G., Beckage, B., 2010. Projecting the distribution of forests in New England in response to climate change. Divers. Distrib. 16, 144–158. doi:10.1111/j.1472-4642.2009.00628.x

The Wilderness Act, 1964. . 88th Congress of the United States of America.

Torrence, C., Compo, G.P., 1995. A Practical Guide to Wavelet Analysis.

Tryhorn, L., DeGaetano, A., 2011. A comparison of techniques for downscaling extreme precipitation over the Northeastern United States. Int. J. Climatol. 31, 1975–1989. doi:10.1002/joc.2208

USDA, 2008. White Mountain National Forest Monitoring and Evaluation Report.

USDA, 2009. Green Mountain National Forest Annual Monitoring and Evaluation Report.

Venegas, S.A., Mysak, L.A., Straub, D.N., 1996. Atmosphere-ocean coupled variability in the south atlantic. Montreal.

Wang, J., Wang, C.-H., Chen, C.L.P., 2014. The Bounded Capacity of Fuzzy Neural Networks (FNNs) via a New Fully Connected Neural Fuzzy Inference System (F-CONFIS) with Its Applications. IEEE Trans. Fuzzy Syst. 1–1. doi:10.1109/TFUZZ.2013.2292972

Wang, W., Ding, J., 2003. Wavelet Network Model and Its Application to the Prediction of Hydrology. Science (80-. ). 1.

Wilby, R.L., Dawson, C.W., 2007. SDSM 4.2 - A decision support tool for the assessment of regional climate change impacts.

Wunsch, C., 1999. The Interpretation of Short Climate Records, with Comments on the North Atlantic and Southern Oscillations. Bull. Am. Meteorol. Soc. 80, 245–255. doi:10.1175/1520-0477(1999)080<0245:TIOSCR>2.0.CO;2

Young, P.C., Ng, C.N., Lane, K., Parker, D., 1991. Recursive Forecasting, Smoothing and Seasonal Adjustment of Nonstationary Environmental Data. J. Forecast. 10, 57–89. doi:10.1002/for.3980100105

Zhou, L., 2003. Relation between interannual variations in satellite measures of northern forest greenness and climate between 1982 and 1999. J. Geophys. Res. 108, 4004. doi:10.1029/2002JD002510